GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 27, 2003, 12:33:52 ; Search time 35 Seconds
(without alignments)
921.333 Million cell updates/sec

Title: US-09-981-087A-2
Perfect score: 1218
Sequence: 1 MGRGRVQLKRIENKINRQVT.....LTEPNSLLPAWMLRPTTTNE 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1218	100.0	242	20	AAW80998	Arabidopsis AGL8-l
2	1218	100.0	242	20	AAW98921	Arabidopsis AGL8 p
3	1218	100.0	242	21	AAG28304	Arabidopsis thalia
4	909	74.6	180	21	AAG28305	Arabidopsis thalia
5	753	61.8	246	20	AAW95018	Petunia MADS box t
6	742.5	61.0	242	21	AAB32574	Eucalyptus grandis
7	734.5	60.3	244	17	AAR99634	Eucalyptus SQE-1 p
8	732.5	60.1	245	17	AAR99636	Eucalyptus SQE-2L
9	707.5	58.1	205	17	AAR99635	Eucalyptus SQE-2S
10	669.5	55.0	262	21	AAY69920	MADS box protein.
11	662.5	54.4	255	18	AAW23811	Arabidopsis floral
12	662.5	54.4	255	20	AAW67571	Arabidopsis LEAFY
13	662.5	54.4	256	19	AAW43109	APETALA1 gene prod
14	662.5	54.4	256	19	AAW39131	Arabidopsis floral
15	662.5	54.4	256	19	AAW43326	Arabidopsis floral
16	662.5	54.4	256	21	AAG52893	Arabidopsis thalia
17	662.5	54.4	256	21	AAY84912	Amino acid sequenc
18	662.5	54.4	256	21	AAY67550	A. thaliana AP1 pr
19	662.5	54.4	256	21	AAY78880	APETALA1 (AP1) nuc
20	662.5	54.4	300	21	AAB19241	Amino acid sequenc
21	662.5	54.4	302	19	AAW69327	Arabidopsis thalia
22	661.5	54.3	256	21	AAG17300	Arabidopsis thalia
23	660	54.2	245	22	AAB50265	Strawberry floweri
24	654	53.7	253	19	AAW69329	Brassica oleracea
25	654	53.7	253	19	AAW43110	APETALA1 gene prod
26	654	53.7	253	19	AAW39132	Brassica floral me
27	654	53.7	253	19	AAW43327	Brassica floral me
28	654	53.7	253	21	AAB19242	Amino acid sequenc
29	654	53.7	253	21	AAY67551	B. oleracea AP1 pr
30	654	53.7	253	21	AAY78881	Brassica oleracea
31	653	53.6	130	21	AAG28306	Arabidopsis thalia
32	648	53.2	253	18	AAW23812	Brassica oleracea
33	647	53.1	255	18	AAW23813	Cauliflower floral
34	647	53.1	255	19	AAW69330	Brassica oleracea
35	647	53.1	255	19	AAW43111	APETALA1 gene prod
36	647	53.1	255	19	AAW43328	Cauliflower floral
37	647	53.1	255	19	AAW39133	Cauliflower floral
38	647	53.1	255	21	AAB19243	Amino acid sequenc
39	647	53.1	255	21	AAY67552	B. oleracea var. b
40	647	53.1	255	21	AAY78882	Cauliflower APETAL
41	627	51.5	255	18	AAW23815	Arabidopsis floral
42	627	51.5	255	19	AAW69324	Arabidopsis thalia
43	627	51.5	255	19	AAW43113	CAULIFLOWER gene p
44	627	51.5	255	19	AAW39135	Arabidopsis floral
45	627	51.5	255	19	AAW43330	Arabidopsis floral

ALIGNMENTS

RESULT 1
AAW80998
ID AAW80998 standard; Protein; 242 AA.
XX
AC AAW80998;
XX
DT 10-MAY-1999 (first entry)
XX
DE Arabidopsis AGL8-like gene product.
XX
KW AGL8-like gene product; agamous-like 8 gene; transcription factor;
KW seed dispersal; dehiscence; transgenic plant.
XX
OS Arabidopsis thaliana.
XX
PN WO9900502-A1.
XX
PD 07-JAN-1999.
XX
PF 25-JUN-1998; 98WO-US13208.
XX
PR 28-APR-1998; 98US-0067800.
PR 27-JUN-1997; 97US-0051030.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Ferrandiz C, Yanofsky MF;
XX
DR WPI; 1999-095747/08.
DR N-PSDB; AAV99856.
XX
PT Use of agamous-like nucleic acids - useful for the production of
PT transgenic seed plants in which dehiscence is modified resulting in

PT delayed seed dispersal
XX
PS Claim 3; Fig 6; 126pp; English.
XX
CC This is the deduced amino acid sequence of the Arabidopsis
CC agamous-like 8 (AGL8) gene product. AGL8 is a transcription factor
CC that regulates dehiscence. The invention provides a transgenic
CC seed plant that is characterised by delayed seed dispersal due to
CC ectopic expression of a nucleic acid molecule (see AAV99856) encoding
CC an AGL8-like gene product. The ectopically expressed nucleic acid
CC molecule may be operatively linked to an exogenous regulatory
CC element such as a dehiscence zone-selective regulatory element
CC derived from the Arabidopsis AGL1 gene (see AAV99857) or AGL5 gene
CC (see AAV99858). The transgenic seed plants include members of the
CC Brassicaceae, such as rapeseed, and members of the Fabaceae, such
CC as soybean, pea, lentil and bean. A plant expression vector
CC comprising a dehiscence zone-selective regulatory element, and a
CC kit for producing the transgenic seed plants are also provided.
XX
SQ Sequence 242 AA;

Query Match 100.0%; Score 1218; DB 20; Length 242;
Best Local Similarity 100.0%; Pred. No. 1.1e-99;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy    241 NE 242
          ||
Db    241 NE 242
```

RESULT 2
AAW98921
ID AAW98921 standard; Protein; 242 AA.
XX
AC AAW98921;
XX
DT 06-MAY-1999 (first entry)
XX
DE Arabidopsis AGL8 protein sequence.
XX
KW Arabidopsis; AGL8; agamous-like 8; seed plant; fruit size.
XX
OS Arabidopsis sp.
XX
PN WO9900503-A1.
XX
PD 07-JAN-1999.
XX
PF 26-JUN-1998; 98WO-US13249.
XX
PR 27-JUN-1997; 97US-0051030.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
PA (REGC) UNIV CALIFORNIA.
XX
PI Ferrandiz C, Gu Q, Martienssen R, Yanofsky MF;
XX

DR WPI; 1999-095748/08.
DR N-PSDB; AAX18594.
XX
PT Use of agamous-like 8 related nucleic acids - useful for producing
PT transgenic seed plants which can produce seed or fruit of increased
PT or decreased size
XX
PS Claim 3; Fig 1; 68pp; English.
XX
CC The present invention describes a non-naturally occurring seed plant
CC comprising an ectopically expressed nucleic acid molecule encoding an
CC agamous-like 8 (AGL8)-related gene product. The seed plant characterised
CC by producing seeds of increased size. The present sequence represents
CC Arabidopsis AGL-8. The seed plant can be used for producing seed or
CC fruit of increased or decreased size. The present invention can relate
CC to seed plants such as, e.g. canola, soybean, pea, broad bean, peanut,
CC coconut, sunflower, cotton, coffee, cacao, citrus, grape, apples,
CC strawberry, olive and tomato.
XX
SQ Sequence 242 AA;

Query Match 100.0%; Score 1218; DB 20; Length 242;
Best Local Similarity 100.0%; Pred. No. 1.1e-99;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy    241 NE 242
          ||
Db    241 NE 242
```

RESULT 3
AAG28304
ID AAG28304 standard; Protein; 242 AA.
XX
AC AAG28304;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33471.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.

PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 100.0%; Score 1218; DB 21; Length 242;
Best Local Similarity 100.0%; Pred. No. 1.1e-99;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy    241 NE 242
          ||
Db    241 NE 242
```

RESULT 4
AAG28305
ID AAG28305 standard; Protein; 180 AA.
XX
AC AAG28305;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 33472.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX

PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 74.6%; Score 909; DB 21; Length 180;
Best Local Similarity 100.0%; Pred. No. 1.6e-72;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      63 MERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSLSL 122
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSLSL 60

Qy     123 KELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTGQQ 182
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 KELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTGQQ 120

Qy     183 EGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTTNE 242
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 EGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTTNE 180
```

RESULT 5
AAW95018
ID AAW95018 standard; Protein; 246 AA.
XX
AC AAW95018;
XX
DT 21-MAY-1999 (first entry)
XX
DE Petunia MADS box transcription factor, FBP10.
XX
KW Petunia; MADS box; transcription factor; FBP10; flowering; inhibition;
KW inflorescence; vegetative organ; storage organ; wood production;
KW nitrogen fixation; ornamental plant.
XX
OS Petunia sp.
XX
FH Key Location/Qualifiers

```
FT   Region          1..56
FT                   /note= "MADS box region"
XX
PN   WO9904003-A1.
XX
PD   28-JAN-1999.
XX
PF   18-JUL-1997;   97WO-NL00424.
XX
PR   18-JUL-1997;   97WO-NL00424.
XX
PA   (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
XX
PI   Angenent GC,  Busscher M,  Franken J;
XX
DR   WPI; 1999-132247/11.
DR   N-PSDB; AAX17796.
XX
PT   New DNA sequence encoding a MADS box transcription factor, FBP10 -
PT   useful for producing transgenic plants in which flowering is
PT   inhibited
XX
PS   Claim 1; Fig 1; 38pp; English.
XX
CC   This represents a petunia MADS box transcription factor, FBP10 protein.
CC   This can be used in a process of producing a transgenic plant in which
CC   flowering is inhibited due to the expression of endogenous FBP10 or a
CC   homologous gene being inhibited; This is useful for producing plants
CC   without inflorescences or flowers which is beneficial for directing the
CC   high energy, which seeds and flowers utilise, to form vegetative organs
CC   or deposition in storage organs. Therefore, the total biomass of
CC   vegetative tissues such as roots, tubers, stem and leaves can be
CC   increased. This is particularly useful for plants used economically e.g.
CC   lettuce, spinach, chicory, sugar beet, potato, trees for wood production,
CC   tobacco, grasses, plants for nitrogen fixation, ornamental plants for
CC   production of cuttings.
XX
SQ   Sequence   246 AA;

  Query Match             61.8%;  Score 753;  DB 20;  Length 246;
  Best Local Similarity   63.1%;  Pred. No. 1.5e-58;
  Matches  157;  Conservative   34;  Mismatches   38;  Indels   20;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||:|||||||||||||||||||||||||||||||||||| |||||:|||||||:||
Db       1 MGRGRVQMKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| |:::||| | |   :| |||||||||:||:::|:|::|||||||||
Db      61 SCMERILERYERYSYAERQLVSTDHSSPGSWNLEHAKLKARIEVVQRNQRHYMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |:|:||:|| |||:::| ||||||| | |||| ||||||:||: || | ||:|||||:
Db     121 SMKDLQNLEQQLDSSLKHIRSRKNQLMHESISELQKKDKSLQEQNNLLSKKVKEREKELA 180

Qy     181 Q-----QEGQLVQCSNSSSVLLPQ--------YCVTSSRDGFVERVGGENGGASSLTEPN 227
           |     |:    : ::||| :|||          |: |       ||  |||   :|
Db     181 QQTQWEQQNNHHEINSSSSFVLPQPLDSPHLGEAYQSTVD------NGEVEGASQ-QQPA 233

Qy     228 SLLPAWMLR 236
           : :| ||||
Db     234 NTMPPWMLR 242


RESULT 6
AAB32574
ID   AAB32574 standard; Protein; 242 AA.
XX
AC   AAB32574;
XX
DT   25-JAN-2001  (first entry)
XX
DE   Eucalyptus grandis transcription factor protein sequence #32.
XX
KW   Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW   poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW   basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW   homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW   type 2 Cys2His2; CCAAT box element; MYB.
XX
OS   Eucalyptus grandis.
XX
PN   WO200053724-A2.
XX
PD   14-SEP-2000.
XX
PF   09-MAR-2000; 2000WO-US06112.
XX
PR   11-MAR-1999;   99US-0266513.
PR   18-AUG-1999;   99US-0149485.
XX
PA   (GENE-) GENESIS RES & DEV CORP LTD.
PA   (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI   Wood M,  McGrath A,  Shenk MA,  Glenn M;
XX
DR   WPI; 2000-579369/54.
XX
PT   New isolated polynucleotide encoding a plant transcription factor for
PT   producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT   having modified gene expression or modified activity of a polypeptide
PT   -
XX
PS   Claim 8; Page 211; 747pp; English.
XX
CC   The present invention relates to novel plant transcription factors from
CC   Eucalyptus grandis or Pinus radiata. The present sequence is one such
CC   transcription factor. The transcription factor may be used to produce a
CC   plant having modified gene expression such as a woody plant e.g. a
CC   eucalyptus, pine, acacia, poplar, sweetgum, teak, or mahogany species or
CC   to modify the activity of a polypeptide in a plant. The transcription
CC   factors of the present invention are members from the following families
CC   of regulatory proteins: bZIP, bZIP family of G-box binding factors, basic
CC   helix-loop-helix zipper, homeotic/homeodomain/homeobox/MADS, homeodomain
CC   zipper, LIM domain, AP2 and EREBs, zinc finger domains of type 2
CC   Cys2His2, CCAAT box elements and MYB.
XX
SQ   Sequence   242 AA;

  Query Match             61.0%;  Score 742.5;  DB 21;  Length 242;
  Best Local Similarity   64.6%;  Pred. No. 1.3e-57;
  Matches  155;  Conservative   34;  Mismatches   44;  Indels    7;  Gaps    5;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||:|||||||||||||||||:|||||||||||||||||||:||:|||||||||||
Db       1 MGRGRLQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALIIFSAKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| ||: |::  :    :| |||||||||:||| :| |:||||||||||
Db      61 SCMERILERYERYSYSEHQVLASETESIGSWTLEHAKLKARLEVLHRNYRHFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |||:||:|| ||::|:| ||||||| | |||| |||||:|||: || | :|:||:|:
Db     121 SLKDLQNLEQQLESALKHIRSRKNQLMHESISVLQKKDRALQEQNNLLTRKVKEKERALA 180

Qy     181 QQEGQLVQCSNS--SSVLLPQYCVTSSRDG-FVERVGGENGGASSLTEP--NSLLPAWML 235
           ||  |  |  ::  | |:|| |  :   :| :  |  | : |  :||:|   :||| |||
Db     181 QQ-AQWEQQDHALDSPVVLPHYLPSLDINGSYQARHNGHDDG-ENLTQPRAGTLLPPWML 238


RESULT 7
AAR99634
ID   AAR99634 standard; Protein; 244 AA.
XX
AC   AAR99634;
XX
```

```
DT   01-OCT-1996  (first entry)
XX
DE   Eucalyptus SQE-1 protein.
XX
KW   Eucalyptus; reproductive gene; SQE-1; Arabidopsis; agomous gene;
KW   Antirrhinum; plena gene; sterility; fertility; forest tree;
KW   transgenic plant; flower development; antisense; ribozyme.
XX
OS   Eucalyptus globulus.
XX
FH   Key             Location/Qualifiers
FT   Region          2..57
FT                   /label= MADS_box
XX
PN   AU9539013-A.
XX
PD   30-MAY-1996.
XX
PF   22-NOV-1995;   95AU-0039013.
XX
PR   22-NOV-1994;   94AU-0009589.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX
PI   Dennis ES,  Harcourt RL,  Kyozuka J,  Llewellyn D;
PI   Peacock WJ,  Southerton S;
XX
DR   WPI; 1996-278411/29.
DR   N-PSDB; AAT34429.
XX
PT   Eucalyptus reproductive genes - useful for prodn. of sterile
PT   Eucalyptus trees useful for establishing wood lot plantations or in
PT   re-forestation projects
XX
PS   Disclosure; Page 31-32; 60pp; English.
XX
CC   SQE-1 protein (AAR99634) is the product of a cDNA clone, SQE1
CC   (AAT34429), isolated from Eucalyptus globulus cDNA on the basis of
CC   its homology to the Arabidopsis agamous (AG) and Antirrhinum plena
CC   (PLE) flower development gene MADS box regions.  Related proteins,
CC   SQE-2S (AAR99635) and SQE-2L (AAR99636), have also been identified.
CC   Antisense or ribozyme constructs of SQE, or of FLE2 and AGE genes
CC   (see also AAT34426, AAT34428 and AAT34432), may be useful in producing
CC   sterile transgenic Eucalyptus trees by minimising inflorescence.
XX
SQ   Sequence   244 AA;

  Query Match             60.3%;  Score 734.5;  DB 17;  Length 244;
  Best Local Similarity   64.4%;  Pred. No. 6.6e-57;
  Matches  154;  Conservative   34;  Mismatches   44;  Indels    7;  Gaps    5;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||:|||||||||||||||||:||||||||||||||||||||||:||||||||||
Db       1 MGRGRLQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALIVFSAKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| |:: |::  :    :| |||||||||:||| :| |:||||||||
Db      61 SCMERILERYERYSYAEHQVLASETESIGSWTLEHAKLKARLEVLHRNYRHFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |||:||:|| ||::|:| ||||||| | |||| :||||:|||: || | |||||:|:
Db     121 SLKDLQNLEQQLESALKHIRSRKNQLMHESISVIQKKDRALQEQNNLLTKKIKEKERALA 180

Qy     181 QQEGQLVQCSNS--SSVLLPQYCVTSSRDG-FVERVGGENGGASSLTEP--NSLLPAWM 234
           ||  |  |  ::  | |:|| |  :   :| :  |  | : |  :||:|   :||| |:
Db     181 QQ-AQWEQQDHALDSPVVLPHYLPSLDINGSYQARHNGHDDG-ENLTQPRAGTLLPPWI 237


RESULT 8
AAR99636
ID   AAR99636 standard; Protein; 245 AA.
XX
AC   AAR99636;
XX
DT   01-OCT-1996  (first entry)
XX
DE   Eucalyptus SQE-2L protein.
XX
KW   Eucalyptus; reproductive gene; SQE-2L; Arabidopsis; agomous gene;
KW   Antirrhinum; plena gene; sterility; fertility; forest tree;
KW   transgenic plant; flower development; antisense; ribozyme.
XX
OS   Eucalyptus globulus.
XX
FH   Key             Location/Qualifiers
FT   Region          2..57
FT                   /label= MADS_box
XX
PN   AU9539013-A.
XX
PD   30-MAY-1996.
XX
PF   22-NOV-1995;   95AU-0039013.
XX
PR   22-NOV-1994;   94AU-0009589.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX
PI   Dennis ES,  Harcourt RL,  Kyozuka J,  Llewellyn D;
PI   Peacock WJ,  Southerton S;
XX
DR   WPI; 1996-278411/29.
DR   N-PSDB; AAT34431.
XX
PT   Eucalyptus reproductive genes - useful for prodn. of sterile
PT   Eucalyptus trees useful for establishing wood lot plantations or in
PT   re-forestation projects
XX
PS   Disclosure; Page 36-37; 60pp; English.
XX
CC   SQE-2L protein (AAR99636) is the product of a cDNA clone, SQE2L
CC   (AAT34431), isolated from Eucalyptus globulus cDNA on the basis of
CC   its homology to the Arabidopsis agamous (AG) and Antirrhinum plena
CC   (PLE) flower development gene MADS box regions.  Related proteins,
CC   SQE-1 (AAR99634) and SQE-2S (AAR99635), have also been identified.
CC   Antisense or ribozyme constructs of SQE, or of FLE2 and AGE genes
CC   (see also AAT34426, AAT34428 and AAT34432), may be useful in producing
CC   sterile transgenic Eucalyptus trees by minimising inflorescence.
XX
SQ   Sequence   245 AA;

  Query Match             60.1%;  Score 732.5;  DB 17;  Length 245;
  Best Local Similarity   63.4%;  Pred. No. 1e-56;
  Matches  156;  Conservative   27;  Mismatches   46;  Indels   17;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||| ||||||||||||||:|||||||:||||:||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRCGLLKKAHEISVLCDADVALIVFSTKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           |||||||||:|| |:: |::  :    : || |||||||||:|:|:||::| |||:||||
Db      61 CCMERILERYERYSYAESQVLTNNAETNGNWTLEHAKLKARMEILQKNQKNLMGEELDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ||||||:|||||| |:|:|||||  | | |||| ||:|||||: || | ||:||:||
Db     121 SLKELQNLEHQLDTALKNIRSRKIQLMCESISELQRKDKALQEQNNMLAKKVKEKEKALA 180

Qy     181 QQEGQLVQCSN-------SSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPN----SL 229
           ||      |  |        ||||:| |      :  |      | | | |  |  |     :|
Db     181 QQ----TQWDNPQDDGLTSSSVILSQSLQPLNIGGPYHPSGIEEGAA--LGPPQHRNATL 234

Qy     230 LPAWML 235
             |:|||
Db     235 FPSWML 240
```

```
RESULT 9
AAR99635
ID   AAR99635 standard; Protein; 205 AA.
XX
AC   AAR99635;
XX
DT   01-OCT-1996  (first entry)
XX
DE   Eucalyptus SQE-2S protein.
XX
KW   Eucalyptus; reproductive gene; SQE-2S; Arabidopsis; agomous gene;
KW   Antirrhinum; plena gene; sterility; fertility; forest tree;
KW   transgenic plant; flower development; antisense; ribozyme.
XX
OS   Eucalyptus globulus.
XX
FH   Key             Location/Qualifiers
FT   Region          2..57
FT                   /label= MADS_box
XX
PN   AU9539013-A.
XX
PD   30-MAY-1996.
XX
PF   22-NOV-1995;    95AU-0039013.
XX
PR   22-NOV-1994;    94AU-0009589.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX
PI   Dennis ES,  Harcourt RL,  Kyozuka J,  Llewellyn D;
PI   Peacock WJ,  Southerton S;
XX
DR   WPI; 1996-278411/29.
DR   N-PSDB; AAT34430.
XX
PT   Eucalyptus reproductive genes - useful for prodn. of sterile
PT   Eucalyptus trees useful for establishing wood lot plantations or in
PT   re-forestation projects
XX
PS   Disclosure; Page 34-35; 60pp; English.
XX
CC   SQE-2S protein (AAR99635) is the product of a cDNA clone, SQE2S
CC   (AAT34430), isolated from Eucalyptus globulus cDNA on the basis of
CC   its homology to the Arabidopsis agamous (AG) and Antirrhinum plena
CC   (PLE) flower development gene MADS box regions.  Related proteins,
CC   SQE-1 (AAR99634) and SQE-2L (AAR99636), have also been identified.
CC   Antisense or ribozyme constructs of SQE, or of FLE2 and AGE genes
CC   (see also AAT34426, AAT34428 and AAT34432), may be useful in producing
CC   sterile transgenic Eucalyptus trees by minimising inflorescence.
XX
SQ   Sequence   205 AA;

  Query Match             58.1%;  Score 707.5;  DB 17;  Length 205;
  Best Local Similarity   70.2%;  Pred. No. 1.3e-54;
  Matches  144;  Conservative   22;  Mismatches   28;  Indels   11;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||| |||||||||||||:|||||||:|||||||:||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRCGLLKKAHEISVLCDADVALIVFSTKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
            |||||||||||| |:: |::  :   : || |||||||||:|:|:||::| |||:||||
Db      61 CCMERILERYDRYSYAENQVLTNNAETNGNWTLEHAKLKARMEILQKNQKNLMGEELDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |||||:|||||| |:|:||||| | | |||| ||:||||||: || | ||:||:||
Db     121 SLKELQNLEHQLDTALKNIRSRKIQLMCESISELQRKDKALQEQNNMLAKKVKEKEKALA 180

Qy     181 QQEGQLVQCSN-------SSSVLLP 198
           ||    |  |        |||  ||
Db     181 QQ----TQWDNPQDDGLTSSSGSLP 201
```

```
RESULT 10
AAY69920
ID   AAY69920 standard; Protein; 262 AA.
XX
AC   AAY69920;
XX
DT   11-APR-2000  (first entry)
XX
DE   MADS box protein.
XX
KW   MADS box; transcription factor; flower organ differentiation;
KW   flower bud formation.
XX
OS   Gentiana triflora.
XX
PN   JP11346773-A.
XX
PD   21-DEC-1999.
XX
PF   10-JUN-1998;   98JP-0161895.
XX
PR   10-JUN-1998;   98JP-0161895.
XX
PA   (IWAT-) IWATE KEN.
XX
DR   WPI; 2000-109690/10.
DR   N-PSDB; AAZ49161.
XX
PT   MADS box gene coding transcription controlling factor  - useful for
PT   controlling the flower bud-forming period
XX
PS   Claim 1; Page 7-8; 8pp; Japanese.
XX
CC   This sequence is the MADS box gene protein of the invention. The
CC   protein has the transcription activity of a gene participating in the
CC   differentiation of flower organ. The gene controls the flower bud-forming
CC   period when introduced into a plant body. The gene is useful for the
CC   diversification of flowering time and height.
XX
SQ   Sequence   262 AA;

  Query Match             55.0%;  Score 669.5;  DB 21;  Length 262;
  Best Local Similarity   54.3%;  Pred. No. 4.1e-51;
  Matches  144;  Conservative   36;  Mismatches   50;  Indels   35;  Gaps    5;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||:|||||:|||||||||:|||||||||||||||||||||||||||||||||
Db       1 MGRGRVQLRRIENKISRQVTFSKRRNGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           | || |||||:|| | :|:: | :    :|| :|: || :|:|:|:||| |:: ||||: |
Db      61 SSMESILERYERYSYDEKRINGTESQAKDNWSMEYPKLVSRIELLQKNIRHYEGEDLNPL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           :::|||||| ||| |:| ||:||:| | :|:| ||||:|||| || | ||:|| ||
Db     121 NMRELQSLEQQLDTALKKIRTRKSQLMHDSVSELQKKEKALQEQNNLLAKKLKENEKNIT 180

Qy     181 QQEGQ------LVQCSN-----SSSVLLPQ------------YCVTSSRDGFVERVGGE 216
           ||  |      ||  ||     |   | ||             :  ||      | :| :
Db     181 QQRMQSADQHRLVDVSNHPPPPSPPPLFPQSLSIQPPLQQPGIHFQTS------ETIGSD 234

Qy     217 NG-----GASSLTEPNSLLPAWMLR 236
           |      |  :    |: :| ||:|
Db     235 GGDGNQNGTEAAAGTNTHMPPWMVR 259


RESULT 11
AAW23811
ID   AAW23811 standard; Protein; 255 AA.
XX
AC   AAW23811;
XX
```

```
DT   11-MAY-1998  (first entry)
XX
DE   Arabidopsis floral meristem identity gene APETALA (AP1) product.
XX
KW   Floral meristem identity gene; APETALA1; AP1 gene;
KW   flower development; transgenic plant; angiosperm.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9727287-A1.
XX
PD   31-JUL-1997.
XX
PF   26-JAN-1996;   96WO-US01041.
XX
PR   26-JAN-1996;   96WO-US01041.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Yanofsky MF;
XX
DR   WPI; 1997-393675/36.
DR   N-PSDB; AAT76885 and AAT76892.
XX
PT   Cauliflower floral meristem identity genes and related proteins -
PT   used to develop products for converting shoot meristem to floral
PT   meristem and promoting early flowering in an angiosperm
XX
PS   Disclosure; Fig 1A-B; 132pp; English.
XX
CC   This sequence comprises the Arabidopsis thaliana APETALA (AP1)
CC   floral meristem identity gene product that is involved in the
CC   conversion of shoot meristem to floral meristem.  Arabidopsis AP1
CC   cDNA and genomic DNA sequences are provided (see AAT76885 and AAT76892).
CC   Mutation of the AP1 gene results in replacement of a few basal
CC   flowers by inflorescence shoots that are not subtended by flowers.
CC   When AP1 is ecotopically expressed in shoot meristem, the shoot
CC   meristem is converted to floral meristem and early flowering can
CC   occur.  The invention relates to floral meristem identity genes
CC   AP1, LFY and especially CAL (see AAT76885-97 and AAT99437) and gene
CC   products (see AAW23811-18) and their use in converting shoot meristem
CC   to floral meristem and in promoting early flowering in transgenic
CC   plants, especially angiosperms such as cereal plants, leguminous
CC   plants, oilseed plants, trees, fruit-bearing plants or ornamental
CC   flowers.
XX
SQ   Sequence   255 AA;

  Query Match            54.4%;  Score 662.5;  DB 18;  Length 255;
  Best Local Similarity  65.8%;  Pred. No. 1.6e-50;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||:|||||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
           | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:||||||
Db     121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy     180 GQQE--GQLVQCSNSSSVLLPQ 199
             |||   |  |  |    | ||
Db     181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 12
AAW67571
ID   AAW67571 standard; Protein; 255 AA.
XX
AC   AAW67571;
XX
DT   02-MAR-1999  (first entry)
XX
DE   Arabidopsis LEAFY protein.
XX
KW   Genetically modified plant; transgenic plant; flower meristem identity;
KW   LEAFY; APETALA1; genome; flower meristem development; flowering time;
KW   environmental condition; crop production.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 19
FT                   /note= "encoded by GG"
FT   Misc-difference 39
FT                   /note= "encoded by TT"
FT   Misc-difference 59
FT                   /note= "encoded by AT"
FT   Misc-difference 79
FT                   /note= "encoded by CG"
FT   Misc-difference 99
FT                   /note= "encoded by AG"
FT   Misc-difference 119
FT                   /note= "encoded by GA"
FT   Misc-difference 139
FT                   /note= "encoded by AC"
FT   Misc-difference 159
FT                   /note= "encoded by AG"
FT   Misc-difference 179
FT                   /note= "encoded by CT"
FT   Misc-difference 199
FT                   /note= "encoded by CG"
FT   Misc-difference 219
FT                   /note= "encoded by TT"
FT   Misc-difference 239
FT                   /note= "encoded by GA"
XX
PN   US5844119-A.
XX
PD   01-DEC-1998.
XX
PF   21-DEC-1995;   95US-0576156.
XX
PR   21-DEC-1995;   95US-0576156.
PR   21-DEC-1994;   94US-0360336.
XX
PA   (SALK ) SALK INST BIOLOGICAL STUDIES.
XX
PI   Weigel D;
XX
DR   WPI; 1999-044675/04.
DR   N-PSDB; AAV81351.
XX
PT   Genetically modified plant - having in its genome a heterologous
PT   nucleic acid sequence encoding a specified flower meristem identity
PT   protein
XX
PS   Disclosure; Column 19-22; 19pp; English.
XX
CC   The invention relates to the production of a genetically modified plant
CC   having a heterologous nucleic acid sequence encoding a flower meristem
CC   identity protein selected from LEAFY and/or APETALA1 inserted into
CC   its genome, which results in accelerated flower meristem development in
CC   the plant.  This sequence corresponds to the Arabidopsis thaliana LEAFY
CC   protein, the gene for which is used in the method of the invention.
CC   The modified plants have their flowering times controlled or induced,
CC   regardless of locale or the environmental conditions, thus allowing the
CC   production of crops at any given time.
XX
SQ   Sequence   255 AA;

  Query Match            54.4%;  Score 662.5;  DB 20;  Length 255;
```

```
  Best Local Similarity   65.8%;  Pred. No. 1.6e-50;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
           | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:||||||
Db     121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy     180 GQQE--GQLVQCSNSSSVLLPQ 199
             |||   |  |  |    | ||
Db     181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 13
AAW43109
ID   AAW43109 standard; protein; 256 AA.
XX
AC   AAW43109;
XX
DT   18-JUN-1998  (first entry)
XX
DE   APETALA1 gene product from Arabidopsis thaliana.
XX
KW   Transgenic plant; ectopically expressed; meristem gene; APETALA1;
KW   AP1; CAULIFLOWER; CAL; LEAFY; LFY; reproductive development;
KW   angiosperm.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9746077-A1.
XX
PD   11-DEC-1997.
XX
PF   05-JUN-1996;   96WO-US09429.
XX
PR   05-JUN-1996;   96WO-US09429.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Weigel D,  Yanofsky MF;
XX
DR   WPI; 1998-041769/04.
DR   N-PSDB; AAT86628.
XX
PT   Transgenic plant comprising an ectopically expressed floral meristem
PT   gene  - for promoting early reproductive development and controlling
PT   the time of seed-derived crop harvest(s) in e.g. grapes, beans,
PT   corn, wheat, etc
XX
PS   Claim 12; Page 120; 158pp; English.
XX
CC   This sequence represents a floral meristem identity gene product,
CC   APETALA1 (AP1) from Arabidopsis thaliana. The invention relates to
CC   a non-naturally occurring seed plant comprising a first ectopically
CC   expressible nucleic acid encoding a first floral meristem identity gene
CC   product, provided that the first nucleic acid is not ectopically
CC   expressed due to a mutation in an endogenous TERMINAL FLOWER gene.
CC   The invention describes a method of converting shoot meristem to floral
CC   meristem, especially in order to promote early reproductive development,
CC   in an angiosperm, which comprises introducing a first ectopically
CC   expressible nucleic acid molecule encoding a first floral meristem
CC   identity gene product into the angiosperm. Ectopic expression of the
CC   floral meristem gene product in the shoot meristem tissue of the
CC   transgenic plant, allows selection of the time of seed development in
CC   the plant which can be useful for manipulating the time of crop harvest
CC   in seed-derived crops such as grapes, beans, corn, wheat, rice, hop,
CC   etc., or to hasten the breeding of e.g. tree species, for insect or
CC   disease resistance.
XX
SQ   Sequence   256 AA;

  Query Match             54.4%;  Score 662.5;  DB 19;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 1.6e-50;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
           | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:||||||
Db     121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy     180 GQQE--GQLVQCSNSSSVLLPQ 199
             |||   |  |  |    | ||
Db     181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 14
AAW39131
ID   AAW39131 standard; Protein; 256 AA.
XX
AC   AAW39131;
XX
DT   08-JUN-1998  (first entry)
XX
DE   Arabidopsis floral meristem identity gene product APETALA1.
XX
KW   Floral meristem identity gene product; APETALA1; AP1;
KW   transgenic plant; angiosperm; seed development.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9746078-A1.
XX
PD   11-DEC-1997.
XX
PF   05-JUN-1996;   96WO-US09453.
XX
PR   05-JUN-1996;   96WO-US09453.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Yanofsky MF;
XX
DR   WPI; 1998-041770/04.
DR   N-PSDB; AAV02760.
XX
PT   Nucleic acids encoding AP1 floral meristem identity gene product
PT   from cauliflower and maize - useful for promoting early reproductive
PT   development and controlling the time of seed-derived crop harvest in
PT   e.g. grapes, beans, corn, wheat, etc
XX
PS   Disclosure; Page 120-121; 156pp; English.
XX
CC   This protein comprises Arabidopsis thaliana floral meristem
CC   identity gene product APETALA (AP1) that is involved in the
CC   conversion of shoot meristem to floral meristem.  Mutation of the
CC   AP1 gene (see AAV02760) results in replacement of basal flowers by
CC   inflorescence shoots that are not subtended by flowers.  The
CC   invention relates to an expression vector comprising a nucleic
CC   acid sequence encoding a floral meristem gene product operably
CC   linked to a heterologous regulatory element (see AAV02770-75).  The
CC   floral meristem gene product is selected from AP1, LFY (LEAFY) and
CC   CAL (CAULIFLOWER) (see AAW39131-38).  The expression vector is used to
```

```
CC   convert shoot meristem tissue to floral meristem tissue in
CC   transgenic plants, especially angiosperms or gymnosperms, thereby
CC   promoting early reproductive development in these plants.  This can
CC   be used to manipulate the time of crop harvest and to hasten
CC   breeding time.
XX
SQ   Sequence   256 AA;

  Query Match             54.4%;  Score 662.5;  DB 19;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 1.6e-50;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||:|||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
           | ||||:|| ||| |:| ||:|||| |:||: ||||:||:|: |: | |:||||||
Db     121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy     180 GQQE--GQLVQCSNSSSVLLPQ 199
             |||   |  |  |    | ||
Db     181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 15
AAW43326
ID   AAW43326 standard; Protein; 256 AA.
XX
AC   AAW43326;
XX
DT   08-JUN-1998  (first entry)
XX
DE   Arabidopsis floral meristem identity gene product APETALA1.
XX
KW   Floral meristem identity gene; APETALA1; AP1; transgenic plant;
KW   angiosperm; seed development.
XX
OS   Arabidopsis thaliana.
XX
PN   WO9746079-A1.
XX
PD   11-DEC-1997.
XX
PF   04-JUN-1997;   97WO-US09682.
XX
PR   05-JUN-1996;   96US-0659188.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Yanofsky MF;
XX
DR   WPI; 1998-041771/04.
DR   N-PSDB; AAV06018.
XX
PT   Use of floral meristem identity genes to convert shoot meristem to
PT   floral meristem - for promoting early reproductive development in
PT   the plants and control timing of seed-derived crop harvest(s) in
PT   e.g. grapes, beans, corn and wheat
XX
PS   Claim 6; Page 103-105; 147pp; English.
XX
CC   This protein comprises Arabidopsis thaliana floral meristem
CC   identity gene product APETALA (AP1) that is involved in the
CC   conversion of shoot meristem to floral meristem.  Mutation of the
CC   AP1 gene (see AAV06018) results in replacement of basal flowers by
CC   inflorescence shoots that are not subtended by flowers.  The
CC   invention relates to a recombinant nucleic acid molecule that
CC   comprises an inducible regulatory element (see AAV06025-28) operably
CC   linked to a nucleic acid molecule (see AAV06018-24) encoding a floral
CC   meristem identity gene product (see AAW43326-32), especially AP1,
CC   CAULIFLOWER (CAL) or LEAFY (LFY).  The nucleic acids are used to
CC   convert shoot meristem tissue into floral meristem tissue in
CC   transgenic plants, thereby promoting early reproductive development
CC   in these plants.  This can be useful for manipulating the time of
CC   crop harvest in seed-derived crops and to hasten breeding time.
XX
SQ   Sequence   256 AA;

  Query Match             54.4%;  Score 662.5;  DB 19;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 1.6e-50;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||:|||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
           | ||||:|| ||| |:| ||:|||| |:||: ||||:||:|: |: | |:||||||
Db     121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy     180 GQQE--GQLVQCSNSSSVLLPQ 199
             |||   |  |  |    | ||
Db     181 AQQEQWDQQNQGHNMPPPLPPQ 202


Search completed: January 27, 2003, 12:38:04
Job time : 37 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 27, 2003, 12:37:07 ; Search time 15 Seconds
(without alignments)
474.690 Million cell updates/sec

Title: US-09-981-087A-2
Perfect score: 1218
Sequence: 1 MGRGRVQLKRIENKINRQVT.....LTEPNSLLPAWMLRPTTTNE 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1218	100.0	242	4	US-09-067-800-2	Sequence 2, Appli
2	1218	100.0	242	4	US-09-105-652-2	Sequence 2, Appli
3	1218	100.0	242	4	US-09-349-677-2	Sequence 2, Appli
4	662.5	54.4	255	2	US-08-576-156-2	Sequence 2, Appli
5	662.5	54.4	256	3	US-08-659-188-2	Sequence 2, Appli
6	662.5	54.4	256	3	US-08-655-227-2	Sequence 2, Appli
7	662.5	54.4	256	3	US-08-655-241-2	Sequence 2, Appli
8	662.5	54.4	256	4	US-09-398-326-2	Sequence 2, Appli
9	662.5	54.4	300	2	US-08-592-214A-2	Sequence 2, Appli
10	662.5	54.4	300	3	US-09-149-976-2	Sequence 2, Appli
11	654	53.7	253	2	US-08-592-214A-4	Sequence 4, Appli
12	654	53.7	253	3	US-08-659-188-4	Sequence 4, Appli
13	654	53.7	253	3	US-08-655-227-4	Sequence 4, Appli
14	654	53.7	253	3	US-08-655-241-4	Sequence 4, Appli
15	654	53.7	253	3	US-09-149-976-4	Sequence 4, Appli
16	654	53.7	253	4	US-09-398-326-4	Sequence 4, Appli
17	647	53.1	255	2	US-08-592-214A-6	Sequence 6, Appli
18	647	53.1	255	3	US-08-659-188-6	Sequence 6, Appli
19	647	53.1	255	3	US-08-655-227-6	Sequence 6, Appli
20	647	53.1	255	3	US-08-655-241-6	Sequence 6, Appli
21	647	53.1	255	3	US-09-149-976-6	Sequence 6, Appli
22	647	53.1	255	4	US-09-398-326-6	Sequence 6, Appli
23	627	51.5	255	2	US-08-592-214A-10	Sequence 10, Appl
24	627	51.5	255	3	US-08-659-188-10	Sequence 10, Appl
25	627	51.5	255	3	US-08-655-227-10	Sequence 10, Appl
26	627	51.5	255	3	US-08-655-241-10	Sequence 10, Appl
27	627	51.5	255	3	US-09-149-976-10	Sequence 10, Appl
28	627	51.5	255	4	US-09-398-326-10	Sequence 10, Appl
29	621.5	51.0	251	2	US-08-592-214A-12	Sequence 12, Appl
30	621.5	51.0	251	3	US-08-659-188-12	Sequence 12, Appl
31	621.5	51.0	251	3	US-08-655-227-12	Sequence 12, Appl
32	621.5	51.0	251	3	US-08-655-241-12	Sequence 12, Appl
33	621.5	51.0	251	3	US-09-149-976-12	Sequence 12, Appl
34	621.5	51.0	251	4	US-09-398-326-12	Sequence 12, Appl
35	610.5	50.1	273	2	US-08-592-214A-8	Sequence 8, Appli
36	610.5	50.1	273	3	US-08-659-188-8	Sequence 8, Appli
37	610.5	50.1	273	3	US-08-655-227-8	Sequence 8, Appli
38	610.5	50.1	273	3	US-08-655-241-8	Sequence 8, Appli
39	610.5	50.1	273	3	US-09-149-976-8	Sequence 8, Appli
40	610.5	50.1	273	4	US-09-398-326-8	Sequence 8, Appli
41	548	45.0	150	2	US-08-592-214A-14	Sequence 14, Appl
42	548	45.0	150	3	US-08-659-188-14	Sequence 14, Appl
43	548	45.0	150	3	US-08-655-227-14	Sequence 14, Appl
44	548	45.0	150	3	US-08-655-241-14	Sequence 14, Appl
45	548	45.0	150	3	US-09-149-976-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1
US-09-067-800-2
; Sequence 2, Application US/09067800
; Patent No. 6198024
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Ferrandiz, Cristina
; TITLE OF INVENTION: Seed Plants Characterized by Delayed
; TITLE OF INVENTION: Seed Dispersal
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/067,800
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Campbell, Cathryn A.
; REGISTRATION NUMBER: 31,815
; REFERENCE/DOCKET NUMBER: P-UD 2948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 535-9001
; TELEFAX: (619) 535-8949
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 242 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-067-800-2

Query Match 100.0%; Score 1218; DB 4; Length 242;
Best Local Similarity 100.0%; Pred. No. 4.2e-107;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy     241 NE 242
          ||
Db     241 NE 242


RESULT 2
US-09-105-652-2
; Sequence 2, Application US/09105652
; Patent No. 6229068
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT: Martienssen, Robert
;  APPLICANT: Ferrandiz, Cristina
;  APPLICANT: Gu, Qing
;  TITLE OF INVENTION: Method of Increasing Fruit Size in a Plant
;  FILE REFERENCE: P-UD 3040
;  CURRENT APPLICATION NUMBER: US/09/105,652
;  CURRENT FILING DATE: 1998-06-26
;  EARLIER APPLICATION NUMBER: 60/051,030
;  EARLIER FILING DATE: 1997-06-27
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Arabidopsis sp.
US-09-105-652-2

  Query Match             100.0%;  Score 1218;  DB 4;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-107;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy     241 NE 242
          ||
Db     241 NE 242


RESULT 3
US-09-349-677-2
; Sequence 2, Application US/09349677
; Patent No. 6288305
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Ferrandiz, Cristina
;    TITLE OF INVENTION:  Seed Plants Characterized by Delayed
;    TITLE OF INVENTION:  Seed Dispersal
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/349,677
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/067,800
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 2948
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  242 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-349-677-2

  Query Match             100.0%;  Score 1218;  DB 4;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-107;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy     241 NE 242
          ||
Db     241 NE 242


RESULT 4
US-08-576-156-2
; Sequence 2, Application US/08576156
; Patent No. 5844119
;  GENERAL INFORMATION:
;    APPLICANT:  Weigel, Detlef
;    TITLE OF INVENTION:  Genetically Modified Plants Having Modulated
;    TITLE OF INVENTION:  Flower Development
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
```

```
;       STATE:  CA
;       COUNTRY:  US
;       ZIP:  92037
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  Windows95
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/576,156
;       FILING DATE:  12-21-95
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/360,336
;       FILING DATE:  12-21-94
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Haile, Lisa A.
;       REGISTRATION NUMBER:  38,347
;       REFERENCE/DOCKET NUMBER:  07251/013001
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  619-678-5070
;       TELEFAX:  619-678-5099
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  255 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-576-156-2

  Query Match             54.4%;  Score 662.5;  DB 2;  Length 255;
  Best Local Similarity   65.8%;  Pred. No. 9.6e-55;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
           | ||||:|| ||| |:| ||:|||| |:||: ||||:||:|: |: | |:||||||
Db     121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy     180 GQQE--GQLVQCSNSSSVLLPQ 199
             |||   |  |  |    | ||
Db     181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 5
US-08-659-188-2
; Sequence 2, Application US/08659188
; Patent No. 6002069
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Seed Plants Exhibiting Inducible Early
;    TITLE OF INVENTION:  Reproductive Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,188
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1946
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  256 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-659-188-2

  Query Match             54.4%;  Score 662.5;  DB 3;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 9.7e-55;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
           | ||||:|| ||| |:| ||:|||| |:||: ||||:||:|: |: | |:||||||
Db     121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy     180 GQQE--GQLVQCSNSSSVLLPQ 199
             |||   |  |  |    | ||
Db     181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 6
US-08-655-227-2
; Sequence 2, Application US/08655227
; Patent No. 6025483
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Maize and Cauliflower APETALA1 Gene
;    TITLE OF INVENTION:  Products and Nucleic Acid Molecules Encoding Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,227
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 2143
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  256 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-655-227-2

  Query Match             54.4%;  Score 662.5;  DB 3;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 9.7e-55;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
          | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:||||||
Db    121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy    180 GQQE--GQLVQCSNSSSVLLPQ 199
            |||   |  |  |    | ||
Db    181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 7
US-08-655-241-2
; Sequence 2, Application US/08655241
; Patent No. 6025543
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Weigel, Detlef
;    TITLE OF INVENTION:  Seed Plants Exhibiting Early Reproductive
;    TITLE OF INVENTION:  Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,241
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  CLASS 800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1894
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  256 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-655-241-2

  Query Match             54.4%;  Score 662.5;  DB 3;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 9.7e-55;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
          | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:||||||
Db    121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy    180 GQQE--GQLVQCSNSSSVLLPQ 199
            |||   |  |  |    | ||
Db    181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 8
US-09-398-326-2
; Sequence 2, Application US/09398326
; Patent No. 6355863
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Seed Plants Exhibiting Inducible Early
;    TITLE OF INVENTION:  Reproductive Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/398,326
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/659,188
;      FILING DATE:  05-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 3739
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  256 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-398-326-2

  Query Match             54.4%;  Score 662.5;  DB 4;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 9.7e-55;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
          | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:||||||
```

```
Db    121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy    180 GQQE--GQLVQCSNSSSVLLPQ 199
          |||   |  |  |    | ||
Db    181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 9
US-08-592-214A-2
; Sequence 2, Application US/08592214A
; Patent No. 5811536
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identify
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,214A
;      FILING DATE:  26-JAN-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1927
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  300 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-592-214A-2

  Query Match             54.4%;  Score 662.5;  DB 2;  Length 300;
  Best Local Similarity   65.8%;  Pred. No. 1.2e-54;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
          | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:|||||
Db    121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy    180 GQQE--GQLVQCSNSSSVLLPQ 199
          |||   |  |  |    | ||
Db    181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 10
US-09-149-976-2
; Sequence 2, Application US/09149976
; Patent No. 6127123
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identity
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/149,976
;      FILING DATE:  09-SEP-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/592,214
;      FILING DATE:  26-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 3291
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  300 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-149-976-2

  Query Match             54.4%;  Score 662.5;  DB 3;  Length 300;
  Best Local Similarity   65.8%;  Pred. No. 1.2e-54;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
          | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:|||||
Db    121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy    180 GQQE--GQLVQCSNSSSVLLPQ 199
          |||   |  |  |    | ||
Db    181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 11
US-08-592-214A-4
; Sequence 4, Application US/08592214A
; Patent No. 5811536
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identify
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
```

```
;       COUNTRY:  United States
;       ZIP:  92122
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/592,214A
;       FILING DATE:  26-JAN-1996
;       CLASSIFICATION:  536
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Campbell, Cathryn A.
;       REGISTRATION NUMBER:  31,815
;       REFERENCE/DOCKET NUMBER:  P-UD 1927
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 535-9001
;       TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  253 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-592-214A-4

  Query Match             53.7%;  Score 654;  DB 2;  Length 253;
  Best Local Similarity   66.5%;  Pred. No. 6e-54;
  Matches  123;  Conservative   36;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||:||:||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLMKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDSNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           | ||||:|| ||| |:| ||||||| |::||: ||:|:||:|: |: | |:|||||
Db     121 SPKELQNLEQQLDTALKHIRSRKNQLMYDSINELQRKEKAIQEQNSMLSKQIKERENVLR 180

Qy     181 QQEGQ 185
             |: |
Db     181 AQQEQ 185


RESULT 12
US-08-659-188-4
; Sequence 4, Application US/08659188
; Patent No. 6002069
;  GENERAL INFORMATION:
;     APPLICANT:  Yanofsky, Martin F.
;     TITLE OF INVENTION:  Seed Plants Exhibiting Inducible Early
;     TITLE OF INVENTION:  Reproductive Development and Methods of Making Same
;     NUMBER OF SEQUENCES:  26
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Campbell and Flores
;       STREET:  4370 La Jolla Village Drive, Suite 700
;       CITY:  San Diego
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  92122
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/659,188
;       FILING DATE:  05-JUN-1996
;       CLASSIFICATION:  800
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Campbell, Cathryn A.
;       REGISTRATION NUMBER:  31,815
;       REFERENCE/DOCKET NUMBER:  P-UD 1946
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 535-9001
;       TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  253 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
US-08-659-188-4

  Query Match             53.7%;  Score 654;  DB 3;  Length 253;
  Best Local Similarity   66.5%;  Pred. No. 6e-54;
  Matches  123;  Conservative   36;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||:||:||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLMKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDSNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           | ||||:|| ||| |:| ||||||| |::||: ||:|:||:|: |: | |:|||||
Db     121 SPKELQNLEQQLDTALKHIRSRKNQLMYDSINELQRKEKAIQEQNSMLSKQIKERENVLR 180

Qy     181 QQEGQ 185
             |: |
Db     181 AQQEQ 185


RESULT 13
US-08-655-227-4
; Sequence 4, Application US/08655227
; Patent No. 6025483
;  GENERAL INFORMATION:
;     APPLICANT:  Yanofsky, Martin F.
;     TITLE OF INVENTION:  Maize and Cauliflower APETALA1 Gene
;     TITLE OF INVENTION:  Products and Nucleic Acid Molecules Encoding Same
;     NUMBER OF SEQUENCES:  26
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Campbell and Flores
;       STREET:  4370 La Jolla Village Drive, Suite 700
;       CITY:  San Diego
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  92122
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/655,227
;       FILING DATE:  05-JUN-1996
;       CLASSIFICATION:  800
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Campbell, Cathryn A.
;       REGISTRATION NUMBER:  31,815
;       REFERENCE/DOCKET NUMBER:  P-UD 2143
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 535-9001
;       TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  4:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  253 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  protein
```

```
US-08-655-227-4

  Query Match             53.7%;  Score 654;  DB 3;  Length 253;
  Best Local Similarity   66.5%;  Pred. No. 6e-54;
  Matches  123;  Conservative   36;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||:||:||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLMKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDSNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           | ||||:|| ||| |:| |||||||| |::||: ||:|:||:|: |: | |:|||||
Db     121 SPKELQNLEQQLDTALKHIRSRKNQLMYDSINELQRKEKAIQEQNSMLSKQIKERENVLR 180

Qy     181 QQEGQ 185
             |: |
Db     181 AQQEQ 185


RESULT 14
US-08-655-241-4
; Sequence 4, Application US/08655241
; Patent No. 6025543
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Weigel, Detlef
;    TITLE OF INVENTION:  Seed Plants Exhibiting Early Reproductive
;    TITLE OF INVENTION:  Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,241
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  CLASS 800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1894
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  253 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-655-241-4

  Query Match             53.7%;  Score 654;  DB 3;  Length 253;
  Best Local Similarity   66.5%;  Pred. No. 6e-54;
  Matches  123;  Conservative   36;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||:||:||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLMKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDSNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           | ||||:|| ||| |:| |||||||| |::||: ||:|:||:|: |: | |:|||||
Db     121 SPKELQNLEQQLDTALKHIRSRKNQLMYDSINELQRKEKAIQEQNSMLSKQIKERENVLR 180

Qy     181 QQEGQ 185
             |: |
Db     181 AQQEQ 185


RESULT 15
US-09-149-976-4
; Sequence 4, Application US/09149976
; Patent No. 6127123
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identity
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/149,976
;      FILING DATE:  09-SEP-1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/592,214
;      FILING DATE:  26-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 3291
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  4:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  253 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-09-149-976-4

  Query Match             53.7%;  Score 654;  DB 3;  Length 253;
  Best Local Similarity   66.5%;  Pred. No. 6e-54;
  Matches  123;  Conservative   36;  Mismatches   26;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||:||:||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLMKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDSNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           | ||||:|| ||| |:| |||||||| |::||: ||:|:||:|: |: | |:|||||
Db     121 SPKELQNLEQQLDTALKHIRSRKNQLMYDSINELQRKEKAIQEQNSMLSKQIKERENVLR 180

Qy     181 QQEGQ 185
             |: |
```

Db 181 AQQEQ 185

Search completed: January 27, 2003, 12:39:23
Job time : 16 secs

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 27, 2003, 12:39:07 ; Search time 11 Seconds
(without alignments)
443.928 Million cell updates/sec

Title: US-09-981-087A-2
Perfect score: 1218
Sequence: 1 MGRGRVQLKRIENKINRQVT.....LTEPNSLLPAWMLRPTTTNE 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 122226 seqs, 20178551 residues

Total number of hits satisfying chosen parameters: 122226

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1218	100.0	242	9	US-09-978-729A-2	Sequence 2, Appli
2	1218	100.0	242	9	US-09-981-087A-2	Sequence 2, Appli
3	1218	100.0	242	9	US-09-978-382A-2	Sequence 2, Appli
4	1218	100.0	242	9	US-09-978-740A-2	Sequence 2, Appli
5	1218	100.0	242	10	US-09-978-730-2	Sequence 2, Appli
6	662.5	54.4	256	9	US-09-853-450-2	Sequence 2, Appli
7	654	53.7	253	9	US-09-853-450-4	Sequence 4, Appli
8	647	53.1	255	9	US-09-853-450-6	Sequence 6, Appli
9	627	51.5	255	9	US-09-853-450-10	Sequence 10, Appl
10	621.5	51.0	251	9	US-09-853-450-12	Sequence 12, Appl
11	610.5	50.1	273	9	US-09-853-450-8	Sequence 8, Appli
12	548	45.0	150	9	US-09-853-450-14	Sequence 14, Appl
13	484.5	39.8	248	9	US-09-853-450-28	Sequence 28, Appl
14	463.5	38.1	250	9	US-09-853-450-30	Sequence 30, Appl
15	443	36.4	237	9	US-09-853-450-32	Sequence 32, Appl
16	396.5	32.6	214	9	US-09-853-450-34	Sequence 34, Appl
17	390.5	32.1	240	10	US-09-970-624-2	Sequence 2, Appli
18	381	31.3	248	9	US-09-978-729A-4	Sequence 4, Appli
19	381	31.3	248	9	US-09-981-087A-4	Sequence 4, Appli
20	381	31.3	248	9	US-09-978-382A-4	Sequence 4, Appli
21	381	31.3	248	9	US-09-978-740A-4	Sequence 4, Appli
22	381	31.3	248	10	US-09-978-730-4	Sequence 4, Appli
23	378	31.0	246	9	US-09-978-729A-6	Sequence 6, Appli
24	378	31.0	246	9	US-09-981-087A-6	Sequence 6, Appli
25	378	31.0	246	9	US-09-978-382A-6	Sequence 6, Appli
26	378	31.0	246	9	US-09-978-740A-6	Sequence 6, Appli
27	378	31.0	246	10	US-09-978-730-6	Sequence 6, Appli
28	328	26.9	220	9	US-09-853-450-38	Sequence 38, Appl
29	314.5	25.8	210	9	US-09-853-450-36	Sequence 36, Appl
30	310	25.5	173	9	US-09-853-450-40	Sequence 40, Appl
31	285.5	23.4	192	9	US-09-853-450-50	Sequence 50, Appl
32	242	19.9	105	10	US-09-767-215-10	Sequence 10, Appl
33	225	18.5	521	10	US-09-749-728B-19	Sequence 19, Appl
34	221	18.1	365	10	US-09-876-187-4	Sequence 4, Appli
35	221	18.1	365	10	US-09-749-728B-15	Sequence 15, Appl
36	218.5	17.9	507	10	US-09-876-187-2	Sequence 2, Appli
37	218.5	17.9	507	10	US-09-749-728B-13	Sequence 13, Appl
38	215.5	17.7	465	10	US-09-749-728B-17	Sequence 17, Appl
39	215.5	17.7	473	10	US-09-876-187-6	Sequence 6, Appli
40	215.5	17.7	473	10	US-09-876-187-8	Sequence 8, Appli
41	182.5	15.0	676	10	US-09-801-368-302	Sequence 302, App
42	170.5	14.0	452	10	US-09-801-368-336	Sequence 336, App
43	150	12.3	286	10	US-09-801-368-188	Sequence 188, App
44	113	9.3	689	9	US-10-108-605-305	Sequence 305, App
45	112	9.2	2568	10	US-09-866-108-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-978-729A-2
; Sequence 2, Application US/09978729A
; Patent No. US20020178465A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegren, Sarah
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-000950US
; CURRENT APPLICATION NUMBER: US/09/978,729A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-978-729A-2

Query Match 100.0%; Score 1218; DB 9; Length 242;
Best Local Similarity 100.0%; Pred. No. 1.8e-92;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy    241 NE 242
          ||
Db    241 NE 242


RESULT 2
US-09-981-087A-2
; Sequence 2, Application US/09981087A
; Patent No. US20020178466A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Liljegren, Sarah
;  APPLICANT:  Farrandiz, Cristina
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 19452A-000940US
;  CURRENT APPLICATION NUMBER: US/09/981,087A
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: US 60/090,649
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 09/339,998
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Arabidopsis sp.
US-09-981-087A-2

  Query Match             100.0%;  Score 1218;  DB 9;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-92;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy    241 NE 242
          ||
Db    241 NE 242


RESULT 3
US-09-978-382A-2
; Sequence 2, Application US/09978382A
; Publication No. US20020194647A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Liljegren, Sarah
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 19452A-000930US
;  CURRENT APPLICATION NUMBER: US/09/978,382A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 60/090,649
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 09/339,998
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Arabidopsis sp.
US-09-978-382A-2

  Query Match             100.0%;  Score 1218;  DB 9;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-92;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy    241 NE 242
          ||
Db    241 NE 242


RESULT 4
US-09-978-740A-2
; Sequence 2, Application US/09978740A
; Publication No. US20030005481A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Liljegren, Sarah
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 19452A-000960US
;  CURRENT APPLICATION NUMBER: US/09/978,740A
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: US 60/090,649
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 09/339,998
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Arabidopsis sp.
US-09-978-740A-2

  Query Match             100.0%;  Score 1218;  DB 9;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-92;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
```

```
Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy    241 NE 242
          ||
Db    241 NE 242


RESULT 5
US-09-978-730-2
; Sequence 2, Application US/09978730
; Patent No. US20020129403A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Liljegren, Sarah
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 19452A-000920US
;  CURRENT APPLICATION NUMBER: US/09/978,730
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: US 60/090,649
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 09/339,998
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 242
;   TYPE: PRT
;   ORGANISM: Arabidopsis sp.
US-09-978-730-2

  Query Match             100.0%;  Score 1218;  DB 10;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-92;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy    241 NE 242
          ||
Db    241 NE 242


RESULT 6
US-09-853-450-2
; Sequence 2, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 256
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: APETALA1 (AP1)
US-09-853-450-2

  Query Match             54.4%;  Score 662.5;  DB 9;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 4.1e-47;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
          | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:||||||
Db    121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy    180 GQQE--GQLVQCSNSSSVLLPQ 199
               |||   |  |  |    | ||
Db    181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 7
US-09-853-450-4
; Sequence 4, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 253
;   TYPE: PRT
;   ORGANISM: Brassica oleracea
;   FEATURE:
;   OTHER INFORMATION: APETALA1 (AP1)
US-09-853-450-4

  Query Match             53.7%;  Score 654;  DB 9;  Length 253;
  Best Local Similarity   66.5%;  Pred. No. 2e-46;
  Matches  123;  Conservative   36;  Mismatches   26;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||||||||||||||||||||:||:||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLMKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDSNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          | ||||:|| ||| |:| |||||||| |::||: ||:|:||:|: |: | |:||||||
```

```
Db    121 SPKELQNLEQQLDTALKHIRSRKNQLMYDSINELQRKEKAIQEQNSMLSKQIKERENVLR 180

Qy    181 QQEGQ 185
          |: |
Db    181 AQQEQ 185


RESULT 8
US-09-853-450-6
; Sequence 6, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
;   LENGTH: 255
;   TYPE: PRT
;   ORGANISM: Brassica oleracea var. botrytis
;   FEATURE:
;   OTHER INFORMATION: APETALA1 (AP1)
US-09-853-450-6

  Query Match             53.1%;  Score 647;  DB 9;  Length 255;
  Best Local Similarity   65.9%;  Pred. No. 7.6e-46;
  Matches  122;  Conservative   35;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          |||||||||||||||||||||||||:||:||||||||||||||:||| |||||||| ||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLMKKAHEISVLCDAEVALVVFSHKGKLFEYPTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          |||| ||||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db     61 SCMEEILERYERYSYAERQLIAPESDSNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          | ||||:|| ||| |:| ||||||| |::||: ||:|:||:|: |: | |:|||||
Db    121 SPKELQNLEQQLDTALKHIRSRKNQLMYDSINELQRKEKAIQEQNSMLSKQIKERENVLR 180

Qy    181 QQEGQ 185
          |: |
Db    181 AQQEQ 185


RESULT 9
US-09-853-450-10
; Sequence 10, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
;   LENGTH: 255
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: CAULIFLOWER (CAL)
US-09-853-450-10

  Query Match             51.5%;  Score 627;  DB 9;  Length 255;
  Best Local Similarity   53.6%;  Pred. No. 3.2e-44;
  Matches  128;  Conservative   47;  Mismatches   54;  Indels   10;  Gaps    3;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||:||||||||||||||||:|||||| |||||||||:||||| |||||||::
Db      1 MGRGRVELKRIENKINRQVTFSKRRTGLLKKAQEISVLCDAEVSLIVFSHKGKLFEYSSE 60

Qy     61 SCMERILERYDRYLYSDKQLVGRD--VSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLD 118
          ||||::||||:|| |:::||:  |  |:   || :|:::|||::|:||:|:|:::||:|:
Db     61 SCMEKVLERYERYSYAERQLIAPDSHVNAQTNWSMEYSRLKAKIELLERNQRHYLGEELE 120

Qy    119 SLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKK 178
          :|||:||:|| ||: |:| ||||||| | ||:: ||:|:| :|: |: | |:|||||
Db    121 PMSLKDLQNLEQQLETALKHIRSRKNQLMNESLNHLQRKEKEIQEENSMLTKQIKERENI 180

Qy    179 TGQQEGQLVQCSNS-SSVLLPQ-------YCVTSSRDGFVERVGGENGGASSLTEPNSL 229
             :: |  | : |   |  ||       | :      |:   |   |   :    |:|
Db    181 LKTKQTQCEQLNRSVDDVPQPQPFQHPHLYMIAHQTSPFLNMGGLYQGEDQTAMRRNNL 239


RESULT 10
US-09-853-450-12
; Sequence 12, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
;   LENGTH: 251
;   TYPE: PRT
;   ORGANISM: Brassica oleracea
;   FEATURE:
;   OTHER INFORMATION: CAULIFLOWER (CAL)
US-09-853-450-12

  Query Match             51.0%;  Score 621.5;  DB 9;  Length 251;
  Best Local Similarity   60.2%;  Pred. No. 9e-44;
  Matches  121;  Conservative   42;  Mismatches   35;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||::|||||||||||||||||:||||||||:||||||:||||| ||||||||::
Db      1 MGRGRVEMKRIENKINRQVTFSKRRAGLLKKAHEISILCDAEVSLIVFSHKGKLFEYSSE 60

Qy     61 SCMERILERYDRYLYSDKQLVGRD--VSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLD 118
          ||||::|| |:|| |::|||   |  |:   || :|:::|||::|:||:|:|:::||||:
Db     61 SCMEKVLEHYERYSYAEKQLKVPDSHVNAQTNWSVEYSRLKAKIELLERNQRHYLGEDLE 120

Qy    119 SLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKK 178
          |:|:||||:|| ||| ::| ||||||| | ||:: ||:|:| : : |: | |:|:|||
Db    121 SISIKELQNLEQQLDTSLKHIRSRKNQLMHESLNHLQRKEKEILEENSMLAKQIRERESI 180

Qy    179 TGQQEGQLVQCSNSSSVLLPQ 199
             : |  |  | |  : ||
Db    181 LRTHQNQSEQ-QNRSHHVAPQ 200


RESULT 11
US-09-853-450-8
; Sequence 8, Application US/09853450
; Publication No. US20020194645A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 273
;   TYPE: PRT
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: APETALA1 (AP1)
US-09-853-450-8

  Query Match             50.1%;  Score 610.5;  DB 9;  Length 273;
  Best Local Similarity   49.1%;  Pred. No. 7.9e-43;
  Matches  133;  Conservative   40;  Mismatches   59;  Indels   39;  Gaps    4;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||:||||||||||||||||||||:||||||||||||||||||:|||| ||||:||:||
Db      1 MGRGKVQLKRIENKINRQVTFSKRRNGLLKKAHEISVLCDAEVAVIVFSPKGKLYEYATD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          | |:::|||||:|| |::| |:  :      ||  |: ||||::| ::|  :: |||||:||
Db     61 SRMDKILERYERYSYAEKALISAESESEGNWCHEYRKLKAKIETIQKCHKHLMGEDLESL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          : |||| || |||:::| ||||||:  | |||| ||||:::||: | :| |:: ||:|
Db    121 NPKELQQLEQQLDSSLKHIRSRKSHLMAESISELQKKERSLQEENKALQKELAERQKAVA 180

Qy    181 ---QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERV----------------------GGE 216
             ||: | ||    :     |   :||   |: |                        | |
Db    181 SRQQQQQQQVQWDQQTHA---QAQTSSSSSSFMMRQDQQGLPPPHNICFPPLTMGDRGEE 237

Qy    217 NGGASSLTEPNSL------------LPAWML 235
             |::  :   |             || |||
Db    238 LAAAAAAQQQQPLPGQAQPQLRIAGLPPWML 268


RESULT 12
US-09-853-450-14
; Sequence 14, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 150
;   TYPE: PRT
;   ORGANISM: Brassica oleracea var. botrytis
;   FEATURE:
;   OTHER INFORMATION: CAULIFLOWER (CAL)
US-09-853-450-14

  Query Match             45.0%;  Score 548;  DB 9;  Length 150;
  Best Local Similarity   68.5%;  Pred. No. 4.7e-38;
  Matches  102;  Conservative   31;  Mismatches   14;  Indels    2;  Gaps    1;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||::|||||||||||||||:|||||||||:||||||:||||| |||||||::
Db      1 MGRGRVEMKRIENKINRQVTFSKRRAGLLKKAHEISILCDAEVSLIVFSHKGKLFEYSSE 60

Qy     61 SCMERILERYDRYLYSDKQLVGRD--VSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLD 118
          ||||::|||||:|| |::|||   |  |:   || :|:::|||::|: |:|:|:::||||:
Db     61 SCMEKVLERYERYSYAEKQLKAPDSHVNAQTNWSMEYSRLKAKIELWERNQRHYLGEDLE 120

Qy    119 SLSLKELQSLEHQLDAAIKSIRSRKNQAM 147
          |:|:||||:|| ||| ::| ||||||| |
Db    121 SISIKELQNLEQQLDTSLKHIRSRKNQLM 149


RESULT 13
US-09-853-450-28
; Sequence 28, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
;   LENGTH: 248
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: SEPALLATA1 (SEP1)
US-09-853-450-28

  Query Match             39.8%;  Score 484.5;  DB 9;  Length 248;
  Best Local Similarity   44.0%;  Pred. No. 1.3e-32;
  Matches  109;  Conservative   46;  Mismatches   80;  Indels   13;  Gaps    4;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||:|||||||||||||:|||:||||||:|:||||||||||:||::|||:|: :
Db      1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          | | : |:|| :  |   ::  :   : ||   |: ||| | | |:: :|| :|||  |
Db     61 SNMLKTLDRYQKCSYGSIEVNNKPAKELENSYREYLKLKGRYENLQRQQRNLLGEDLGPL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKE----RE 176
          : |||: || ||| ::| :|| | | | : :| || |:: | : | :|  |: :    |
Db    121 NSKELEQLERQLDGSLKQVRSIKTQYMLDQLSDLQNKEQMLLETNRALAMKLDDMIGVRS 180

Qy    177 KKTGQQEG--QLVQCSN----SSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEP---N 227
             |  ||  | |  ::    |  :  |  |  : : |:   |  |   |::  :    |
Db    181 HHMGGWEGGEQNVTYAHHQAQSQGLYQPLECNPTLQMGYDNPVCSEQITATTQAQAQPGN 240

Qy    228 SLLPAWML 235
            :| |||
Db    241 GYIPGWML 248


RESULT 14
US-09-853-450-30
; Sequence 30, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
```

```
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
;   LENGTH: 250
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: SEPALLATA2 (SEP2)
US-09-853-450-30

  Query Match             38.1%;  Score 463.5;  DB 9;  Length 250;
  Best Local Similarity   41.1%;  Pred. No. 7e-31;
  Matches  109;  Conservative   40;  Mismatches   71;  Indels   45;  Gaps    5;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||:|||||||||||||||:|||:|||||||:|:||||||||:|||||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVSLIVFSNRGKLYEFCST 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           | | : |||| :  |   ::  :   : ||   |: ||| | | |:: :|| :|||| |
Db      61 SNMLKTLERYQKCSYGSIEVNNKPAKELENSYREYLKLKGRYENLQRQQRNLLGEDLGPL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKE-----R 175
           : |||: || ||| ::| :|  | | | : :| || |:  | | | :|  |:::
Db     121 NSKELEQLERQLDGSLKQVRCIKTQYMLDQLSDLQGKEHILLDANRALSMKLEDMIGVRH 180

Qy     176 EKKTGQQEG------------------QLVQC------SNSSSVLLPQYCVTSSRDGFV 210
             |  ||                    | ::|        |  |   |  ||
Db     181 HHIGGGWEGGDQQNIAYGHPQAHSQGLYQSLECDPTLQIGYSHPVCSEQMAVT------- 233

Qy     211 ERVGGENGGASSLTEPNSLLPAWML 235
             | |::       : |  :| |||
Db     234 --VQGQS------QQGNGYIPGWML 250


RESULT 15
US-09-853-450-32
; Sequence 32, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
;   LENGTH: 237
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   OTHER INFORMATION: SEPALLATA3 (SEP3)
US-09-853-450-32

  Query Match             36.4%;  Score 443;  DB 9;  Length 237;
  Best Local Similarity   41.9%;  Pred. No. 3.1e-29;
  Matches  106;  Conservative   44;  Mismatches   73;  Indels   30;  Gaps    7;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||:|||||||||||||||:|||:||||||:|:|||||||||||:||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy      61 SCMERILERYDRYLYS-------DKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFM 113
           | | | |||| :  |         :: :  ::|  :    |: ||| | : |:: :|| :
Db      61 SSMLRTLERYQKCNYGAPEPNVPSREALAVELSSQQ----EYLKLKERYDALQRTQRNLL 116

Qy     114 GEDLDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKK-----DKALQDHNNSL 168
           ||||  || |||:||| |||:::| ||: : | | : :: || |        || : |
Db     117 GEDLGPLSTKELESLERQLDSSLKQIRALRTQFMLDQLNDLQSKLADGYQMPLQLNPN-- 174

Qy     169 LKKIKEREKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNS 228
             :|     |:   |  |  :| :   |  |    : |:  :  |   | |     |:
Db     175 ----QEEVDHYGRHHHQ--QQQHSQAFFQPLECEPILQIGYQGQQDGMGAGPS----VNN 224

Qy     229 LLPAWMLRPTTTN 241
             :  |:  |  ||
Db     225 YMLGWL--PYDTN 235


Search completed: January 27, 2003, 12:43:05
Job time : 12 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 27, 2003, 12:34:12 ; Search time 11 Seconds
(without alignments)
912.479 Million cell updates/sec

Title: US-09-981-087A-2
Perfect score: 1218
Sequence: 1 MGRGRVQLKRIENKINRQVT.....LTEPNSLLPAWMLRPTTTNE 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID		Description
1	1218	100.0	242	1	AGL8_ARATH	Q38876	arabidopsis
2	1112.5	91.3	241	1	AGL8_SINAL	Q41274	sinapis alb
3	743	61.0	250	1	AGL8_SOLTU	Q42429	solanum tub
4	731	60.0	250	1	AGL8_SOLCO	O22328	solanum com
5	709.5	58.3	227	1	AGL8_LYCES	Q40170	lycopersico
6	662.5	54.4	256	1	AP1_ARATH	P35631	arabidopsis
7	656	53.9	254	1	AP1_SINAL	Q41276	sinapis alb
8	509	41.8	233	1	CMB1_DIACA	Q39685	dianthus ca
9	484.5	39.8	248	1	AGL2_ARATH	P29382	arabidopsis
10	472	38.8	241	1	AGL9_PETHY	Q03489	petunia hyb
11	463.5	38.1	250	1	AGL4_ARATH	P29384	arabidopsis
12	461	37.8	251	1	AGL9_ARATH	O22456	arabidopsis
13	449.5	36.9	254	1	AGL9_SINAL	O04067	sinapis alb
14	449	36.9	224	1	AGL9_LYCES	Q42464	lycopersico
15	446	36.6	250	1	AGL9_ARADE	Q38694	aranda debo
16	442	36.3	252	1	AGL6_ARATH	P29386	arabidopsis
17	432.5	35.5	258	1	AGL3_ARATH	P29383	arabidopsis
18	427	35.1	244	1	AG13_ARATH	Q38837	arabidopsis
19	416.5	34.2	248	1	AG_TOBAC	Q43585	nicotiana t
20	415	34.1	242	1	AG_PANGI	Q40872	panax ginse
21	412	33.8	242	1	AG_PETHY	Q40885	petunia hyb
22	403	33.1	230	1	AG11_ARATH	Q38836	arabidopsis
23	403	33.1	252	1	AG_BRANA	Q01540	brassica na
24	401.5	33.0	248	1	AG_LYCES	Q40168	lycopersico
25	399	32.8	252	1	AG_ARATH	P17839	arabidopsis
26	396.5	32.6	214	1	SOC1_ARATH	O64645	arabidopsis
27	381	31.3	248	1	AGL1_ARATH	P29381	arabidopsis
28	378	31.0	246	1	AGL5_ARATH	P29385	arabidopsis
29	376.5	30.9	221	1	AG14_ARATH	Q38838	arabidopsis
30	373.5	30.7	228	1	AG1Y_ARATH	Q9szj6	arabidopsis
31	356	29.2	219	1	AG19_ARATH	O82743	arabidopsis
32	354.5	29.1	264	1	AG15_BRANA	Q39295	brassica na
33	344.5	28.3	265	1	JOIN_LYCES	Q9fuy6	lycopersico
34	344	28.2	268	1	AG15_ARATH	Q38847	arabidopsis
35	332.5	27.3	211	1	AG12_ARATH	Q38841	arabidopsis
36	325.5	26.7	240	1	SVP_ARATH	Q9fvc1	arabidopsis
37	325	26.7	232	1	AP3_ARATH	P35632	arabidopsis
38	324	26.6	209	1	GLOB_TOBAC	Q03416	nicotiana t
39	323	26.5	227	1	AG17_ARATH	Q38840	arabidopsis
40	320	26.3	210	1	FBP1_PETHY	Q03488	petunia hyb
41	315.5	25.9	231	1	MAD1_PETHY	Q07472	petunia hyb
42	308.5	25.3	196	1	FLC_ARATH	Q9s7q7	arabidopsis
43	306.5	25.2	227	1	DEFA_ANTMA	P23706	antirrhinum
44	301	24.7	208	1	PIST_ARATH	P48007	arabidopsis
45	299	24.5	178	1	AG31_ARATH	Q9fpn7	arabidopsis

ALIGNMENTS

```
RESULT 1
AGL8_ARATH
ID   AGL8_ARATH     STANDARD;      PRT;   242 AA.
AC   Q38876;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL8 (Floral homeotic protein AGL8)
DE   (FRUITFULL).
GN   AGL8 OR FUL OR AT5G60910 OR MSL3.3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=96093419; PubMed=8535133;
RA   Mandel M.A., Yanofsky M.F.;
RT   "The Arabidopsis AGL8 MADS box gene is expressed in inflorescence
RT   meristems and is negatively regulated by APETALA1.";
RL   Plant Cell 7:1763-1771(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=98162728; PubMed=9501997;
RA   Nakamura Y., Sato S., Kaneko T., Kotani H., Asamizu E., Miyajima N.,
RA   Tabata S.;
RT   "Structural analysis of Arabidopsis thaliana chromosome 5. III.
RT   Sequence features of the regions of 1,191,918 bp covered by seventeen
RT   physically assigned P1 clones.";
RL   DNA Res. 4:401-414(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT   SSP consortium (Salk/Stanford/PGEC).";
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   FUNCTION, AND TISSUE SPECIFICITY.
RC   STRAIN=cv. Landsberg erecta;
RX   MEDLINE=98171466; PubMed=9502732;
RA   Gu Q., Ferrandiz C., Yanofsky M.F., Martienssen R.;
RT   "The FRUITFULL MADS-box gene mediates cell differentiation during
RT   Arabidopsis fruit development.";
RL   Development 125:1509-1517(1998).
CC   -!- FUNCTION: Probable transcription factor required for normal
CC       pattern of cell division, expansion and differentiation during
CC       morphogenesis of the silique.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- TISSUE SPECIFICITY: Vascular tissue of cauline leaves, floral
CC       shoot apex and valves of carpels and fruits.
CC   -!- INDUCTION: Dramatically up-regulated upon the transition from
CC       vegetative to reproductive development.
```

```
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U33473; AAA97403.1; -.
DR   EMBL; AB008269; BAB10640.1; -.
DR   EMBL; AF386929; AAK62374.1; -.
DR   EMBL; AY072463; AAL66878.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03030; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
SQ   SEQUENCE   242 AA;  27536 MW;  516499731EADB2D2 CRC64;

  Query Match             100.0%;  Score 1218;  DB 1;  Length 242;
  Best Local Similarity   100.0%;  Pred. No. 2.8e-74;
  Matches  242;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy     241 NE 242
           ||
Db     241 NE 242


RESULT 2
AGL8_SINAL
ID   AGL8_SINAL     STANDARD;      PRT;   241 AA.
AC   Q41274;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL8 homolog (MADS B).
GN   AGL8.
OS   Sinapis alba (White mustard) (Brassica hirta).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Sinapis.
OX   NCBI_TaxID=3728;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Flower;
RX   MEDLINE=97077349; PubMed=8919916;
RA   Menzel G., Apel K., Melzer S.;
RT   "Identification of two MADS box genes that are expressed in the
RT   apical meristem of the long-day plant Sinapis alba in transition to
RT   flowering.";
RL   Plant J. 9:399-408(1996).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN APICAL MERISTEMS IN TRANSITION
CC       TO FLOWERING.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U25695; AAB41525.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03168; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
SQ   SEQUENCE   241 AA;  27660 MW;  06826CC598EA94A9 CRC64;

  Query Match             91.3%;  Score 1112.5;  DB 1;  Length 241;
  Best Local Similarity   91.3%;  Pred. No. 2.7e-67;
  Matches  221;  Conservative    9;  Mismatches   11;  Indels    1;  Gaps    1;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||||||||||||||||||||::||||||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALVIFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:||||||||||||||||:|||||||||||||||||||||||||||||||||||||
Db      61 SCMEKILERYDRYLYSDKQLVGRDISQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |||||||||||| |||||||||||||||||||||||| |||||:|||||||||| |
Db     121 SLKELQSLEHQLHAAIKSIRSRKNQAMFESISALQKKDKVLQDHNNALLKKIKEREKNTV 180

Qy     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
           || ||:||||:||:| ||||:||||||| || ||| ||||| |||||||||||| ||
Db     181 HQEVQLIQCSNNSSILQPQYCLTSSRDGFVGRVEGENEGASSLAEPNSLLPAWMLRP-TT 239

Qy     241 NE 242
           ||
Db     240 NE 241


RESULT 3
AGL8_SOLTU
ID   AGL8_SOLTU     STANDARD;      PRT;   250 AA.
AC   Q42429;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL8 homolog (POTM1-1).
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Superior;
RX   MEDLINE=96343939; PubMed=8756601;
RA   Kang S.G., Hannapel D.J.;
RT   "A novel MADS-box gene of potato (Solanum tuberosum L.) expressed
RT   during the early stages of tuberization.";
RL   Plant Mol. Biol. 31:379-386(1996).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U23758; AAA92840.1; -.
DR   EMBL; U23757; AAA92839.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03145; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
SQ   SEQUENCE   250 AA;  28922 MW;  EDE37FFFE793DDC4 CRC64;

  Query Match             61.0%;  Score 743;  DB 1;  Length 250;
  Best Local Similarity   61.8%;  Pred. No. 8.8e-43;
  Matches  155;  Conservative   36;  Mismatches   40;  Indels   20;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||||||||||||||||||| |||||:|||||||: |
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYAND 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           |||||:||||:|| ::::|||  | :   :| |||||||||:|||::|:::::||||:||
Db      61 SCMERLLERYERYSFAERQLVPTDHTSPGSWTLEHAKLKARLEVLQRNQKHYVGEDLESL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ::||||:|||||:|:| |||||| | |||| |||:|:|||: || | ||:|||||:
Db     121 NMKELQNLEHQLDSALKHIRSRKNQLMHESISVLQKQDRALQEQNNQLSKKVKEREKEVA 180

Qy     181 QQE--GQLVQCSNSSSVLLPQ----------YCVTSSRDGFVERVGGENGGASSLTE--- 225
           ||    |     |||: :|||          |  |:  |      |   || ||  :
Db     181 QQNQWDQQNHEINSSTFVLPQQLDSPHLGEAYQNTNVVDN-----GEVEGGNSSQQQGAA 235

Qy     226 PNSLLPAWMLR 236
            |:::| ||||
Db     236 NNTVMPQWMLR 246


RESULT 4
AGL8_SOLCO
ID   AGL8_SOLCO     STANDARD;      PRT;   250 AA.
AC   O22328;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL8 homolog.
GN   SCM1.
OS   Solanum commersonii (Commerson's wild potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4109;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Seppanen M.M.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF002666; AAB65161.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
SQ   SEQUENCE   250 AA;  28741 MW;  C9871403C23ED5D5 CRC64;

  Query Match             60.0%;  Score 731;  DB 1;  Length 250;
  Best Local Similarity   62.2%;  Pred. No. 5.5e-42;
  Matches  153;  Conservative   39;  Mismatches   44;  Indels   10;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||||||||||||||||||| |||||:|||||||:||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           |||||:||||:|| :::|||| | :   :| ||:||||||:|||::|:: ::||||:||
Db      61 SCMERLLERYERYSFAEKQLVPTDHTSPGSWTLENAKLKARLEVLQRNEKLYVGEDLESL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ::||||:||||| :|:| |||||| | |||| |||:|:|||: || | ||:|||||:
Db     121 NMKELQNLEHQLASALKHIRSRKNQLMHESISVLQKQDRALQEQNNQLSKKVKEREKEVE 180

Qy     181 QQE--GQLVQCSNSSSVLLPQ-----YCVTSSRDGFVERVGGENGGASSLTE---PNSLL 230
           ||    |     |||: :|||     :   :|::  |   |   || ||  :    |:::
Db     181 QQNQWDQQNHEINSSTFVLPQQLDSPHLGEASQNTNVVDNGEVEGGNSSQXQGAANNTVM 240

Qy     231 PAWMLR 236
           | ||:|
Db     241 PQWMVR 246


RESULT 5
AGL8_LYCES
ID   AGL8_LYCES     STANDARD;      PRT;   227 AA.
AC   Q40170;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
```

```
DE   Agamous-like MADS box protein AGL8 homolog (TM4).
GN   TDR4.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. VFNT Cherry; TISSUE=Flower;
RX   MEDLINE=93251098; PubMed=1688249;
RA   Pnueli L., Abu-Abeid M., Zamir D., Nacken W., Schwarz-Sommer Z.,
RA   Lifschitz E.;
RT   "The MADS box gene family in tomato: temporal expression during
RT   floral development, conserved secondary structures and homology with
RT   homeotic genes from Antirrhinum and Arabidopsis.";
RL   Plant J. 1:255-266(1991).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- TISSUE SPECIFICITY: FLOWER-SPECIFIC.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X60757; CAA43169.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03186; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
SQ   SEQUENCE   227 AA;  26403 MW;  4976195B3BDE53F5 CRC64;

  Query Match             58.3%;  Score 709.5;  DB 1;  Length 227;
  Best Local Similarity   72.3%;  Pred. No. 1.3e-40;
  Matches  141;  Conservative   25;  Mismatches   24;  Indels    5;  Gaps    1;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          |||||||||||||||||||||||||||||||||||||||||| |||||:|||||||: |
Db      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYAND 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||:|| :::|||| | :  :| ||| |||||:|||::|:::::||||:||
Db     61 SCMERILERYERYSFAEKQLVPTDHTSPVSWTLEHRKLKARLEVLQRNQKHYVGEDLESL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          |:||||:||||||:|:| ||||||| | |||| |||||:|||: || | ||:|||||
Db    121 SMKELQNLEHQLDSALKHIRSRKNQLMHESISVLQKKDRALQEQNNQLSKKVKEREKSAQ 180

Qy    181 QQEGQLVQCSNSSSV 195
          |  |      ||||:
Db    181 QISG-----INSSSL 190


RESULT 6
AP1_ARATH
ID   AP1_ARATH      STANDARD;      PRT;   256 AA.
AC   P35631; Q9LQA8;
```

```
DT   01-JUN-1994 (Rel. 29, Created)
DT   15-JUN-2002 (Rel. 41, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Floral homeotic protein APETALA1 (AGL7 protein).
GN   AP1 OR AGL7 OR AT1G69120 OR F4N2.9.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RX   MEDLINE=93063372; PubMed=1359429;
RA   Mandel M.A., Gustafson-Brown C., Savidge B., Yanofsky M.F.;
RT   "Molecular characterization of the Arabidopsis floral homeotic gene
RT   APETALA1.";
RL   Nature 360:273-277(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016719; PubMed=11130712;
RA   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA   White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA   Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:816-820(2000).
RN   [3]
RP   SEQUENCE OF 1-251 FROM N.A.
RC   STRAIN=Various strains;
RX   MEDLINE=21969421; PubMed=11973317;
RA   Olsen K.M., Womack A., Garrett A.R., Suddith J.I., Purugganan M.D.;
RT   "Contrasting evolutionary forces in the Arabidopsis thaliana floral
RT   developmental pathway.";
RL   Genetics 160:1641-1650(2002).
CC   -!- FUNCTION: CONTROLS FLORAL MERISTEM IDENTITY. IS ALSO REQUIRED FOR
CC       NORMAL DEVELOPMENT OF SEPALS AND PETALS. IS REQUIRED FOR THE
CC       TRANSITION OF AN INFLUORESCENCE MERISTEM INTO A FLORAL MERISTEM.
CC       INTERACTS WITH LEAFY.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN YOUNG FLOWER PRIMORDIA, LATER
CC       BECOMES LOCALIZED TO SEPALS AND PETALS.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Z16421; CAA78909.1; -.
DR   EMBL; AC008262; AAF27070.1; -.
DR   EMBL; AF466771; AAM28447.1; -.
DR   EMBL; AF466772; AAM28448.1; -.
DR   EMBL; AF466773; AAM28449.1; -.
```

```
DR   EMBL; AF466774; AAM28450.1; -.
DR   EMBL; AF466775; AAM28451.1; -.
DR   EMBL; AF466776; AAM28452.1; -.
DR   EMBL; AF466777; AAM28453.1; -.
DR   EMBL; AF466778; AAM28454.1; -.
DR   EMBL; AF466779; AAM28455.1; -.
DR   EMBL; AF466780; AAM28456.1; -.
DR   EMBL; AF466781; AAM28457.1; -.
DR   EMBL; AF466782; AAM28458.1; -.
DR   EMBL; AF466783; AAM28459.1; -.
DR   EMBL; AF466784; AAM28460.1; -.
DR   EMBL; AF466785; AAM28461.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T01775; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Activator; Nuclear protein;
KW   Developmental protein; Polymorphism.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
FT   VARIANT      85     85       S -> P (IN STRAIN CV. CHI-1).
FT   VARIANT      93     93       M -> T (IN STRAIN CV. BLA-1).
FT   VARIANT      99     99       K -> R (IN STRAIN CV. CHI-1).
FT   VARIANT     100    100       A -> T (IN STRAIN CV. LER).
FT   VARIANT     120    120       M -> V (IN STRAIN CV. JL-1).
FT   VARIANT     125    125       L -> P (IN STRAIN CV. BLA-1).
FT   VARIANT     166    166       S -> G (IN STRAIN CV. JL-1).
FT   VARIANT     212    212       L -> P (IN STRAIN CV. CHI-1).
FT   VARIANT     233    233       A -> V (IN STRAIN CV. CO-1).
FT   CONFLICT    184    184       E -> G (IN REF. 3; AAM28458).
FT   CONFLICT    236    236       MISSING (IN REF. 1).
SQ   SEQUENCE   256 AA;  30182 MW;  01E26DE18CE3478C CRC64;

  Query Match             54.4%;  Score 662.5;  DB 1;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 2e-37;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||:|||||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
           | ||||:|| ||| |:| ||:|||| |:||: ||||:||:|: |: | |:|||||
Db     121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy     180 GQQE--GQLVQCSNSSSVLLPQ 199
               |||   |  |  |    | ||
Db     181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 7
AP1_SINAL
ID   AP1_SINAL      STANDARD;      PRT;   254 AA.
AC   Q41276;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Floral homeotic protein APETALA1 (MADS C).
GN   AP1.
OS   Sinapis alba (White mustard) (Brassica hirta).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Sinapis.
OX   NCBI_TaxID=3728;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Flower;
RX   MEDLINE=95334514; PubMed=7610185;
RA   Menzel G., Apel K., Melzer S.;
RT   "Isolation and analysis of SaMADS C, the APETALA 1 cDNA homolog from
RT   mustard.";
RL   Plant Physiol. 108:853-854(1995).
CC   -!- FUNCTION: CONTROLS FLORAL MERISTEM IDENTITY. IS ALSO REQUIRED FOR
CC       NORMAL DEVELOPMENT OF SEPALS AND PETALS. IS REQUIRED FOR THE
CC       TRANSITION OF AN INFLUORESCENCE MERISTEM INTO A FLORAL MERISTEM.
CC       INTERACTS WITH LEAFY (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X81480; CAA57233.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03169; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Activator; Nuclear protein;
KW   Developmental protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
SQ   SEQUENCE   254 AA;  29918 MW;  187FEEDE51BEF743 CRC64;

  Query Match             53.9%;  Score 656;  DB 1;  Length 254;
  Best Local Similarity   64.1%;  Pred. No. 5.3e-37;
  Matches  127;  Conservative   38;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||:|||||||||||||||||||:||| ||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|:::|||| ::
Db      61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           | ||||:|| ||| |:| ||||||| | :||: ||:|:||:|: |: | |:||||||
Db     121 SSKELQNLEQQLDTALKHIRSRKNQLMHDSINELQRKEKAIQEQNSMLSKQIKEREKILR 180

Qy     181 QQEGQLVQCSNSSSVLLP 198
               |: |  | ::  ::  |
Db     181 AQQEQWDQQNHGHNMPPP 198


RESULT 8
CMB1_DIACA
ID   CMB1_DIACA     STANDARD;      PRT;   233 AA.
AC   Q39685;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   MADS box protein CMB1.
GN   CMB1.
```

```
OS   Dianthus caryophyllus (Carnation) (Clove pink).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Caryophyllaceae; Dianthus.
OX   NCBI_TaxID=3570;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Scania; TISSUE=Petal;
RA   Baudinette S.C., Savin K.W.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L40404; AAA62761.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03060; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     58       MADS.
FT   DOMAIN       96    168       K-BOX.
SQ   SEQUENCE   233 AA;  26888 MW;  CB8765618466151A CRC64;

  Query Match             41.8%;  Score 509;  DB 1;  Length 233;
  Best Local Similarity   44.3%;  Pred. No. 2.7e-27;
  Matches  109;  Conservative   45;  Mismatches   68;  Indels   24;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||:|||||||||||||||:|||:||||||:|:|||||||||||||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIVFSNRGKLYEFCST 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||| : |||| |  |    :   :   ::|:   |: ||||:|:||::: || :|||  |
Db      61 SCMNKTLERYQRCSYGSLE-TSQPSKETESSYQEYLKLKAKVDVLQRSHRNLLGEDLGEL 119

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKE------ 174
           | |||: |||||| ::: ||| | | | : :: ||||:: | : | :|  |::|
Db     120 STKELEQLEHQLDKSLRQIRSIKTQHMLDQLADLQKKEEMLFESNRALKTKLEESCASFR 179

Qy     175 -----REKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSL 229
                |:   |  |   : |:|:  :             |: |    :    :|    :
Db     180 PNWDVRQPGDGFFEPLPLPCNNNLQI-----------GYNEATQDQMNATTSAQNVHGF 227

Qy     230 LPAWML 235
             |||
Db     228 AQGWML 233


RESULT 9
AGL2_ARATH
ID   AGL2_ARATH     STANDARD;      PRT;   248 AA.
AC   P29382;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL2.
GN   AGL2.
```

```
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91160981; PubMed=1672119;
RA   Ma H., Yanofsky M.F., Meyerowitz E.M.;
RT   "AGL1-AGL6, an Arabidopsis gene family with similarity to floral
RT   homeotic and transcription factor genes.";
RL   Genes Dev. 5:484-495(1991).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY DURING FLOWER DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M55551; AAA32732.1; -.
DR   PIR; B39534; B39534.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03024; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
SQ   SEQUENCE   248 AA;  28454 MW;  8B70DD4512AC906B CRC64;

  Query Match             39.8%;  Score 484.5;  DB 1;  Length 248;
  Best Local Similarity   44.0%;  Pred. No. 1.2e-25;
  Matches  109;  Conservative   46;  Mismatches   80;  Indels   13;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||:|||||||||||||||:|||:||||||:|:|||||||||||:||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           | | : |:|| :  |   ::  :   : ||   |: ||| | | |:: :|| :|||  |
Db      61 SNMLKTLDRYQKCSYGSIEVNNKPAKELENSYREYLKLKGRYENLQRQQRNLLGEDLGPL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKE----RE 176
           : |||: || ||| ::| :|| | | | : :| || |:: | : | :|  |: :    |
Db     121 NSKELEQLERQLDGSLKQVRSIKTQYMLDQLSDLQNKEQMLLETNRALAMKLDDMIGVRS 180

Qy     177 KKTGQQEG--QLVQCSN----SSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEP---N 227
             |  ||  | |  ::    |  :  |  |  : : |:   |  |   |::  :    |
Db     181 HHMGGWEGGEQNVTYAHHQAQSQGLYQPLECNPTLQMGYDNPVCSEQITATTQAQAQPGN 240

Qy     228 SLLPAWML 235
             :| |||
Db     241 GYIPGWML 248


RESULT 10
AGL9_PETHY
ID   AGL9_PETHY     STANDARD;      PRT;   241 AA.
AC   Q03489;
```

```
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1996 (Rel. 33, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL9 homolog (Floral homeotic protein
DE   FBP2) (Floral binding protein 2).
GN   FBP2.
OS   Petunia hybrida (Petunia).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
OX   NCBI_TaxID=4102;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93005737; PubMed=1356537;
RA   Angenent G.C., Busscher M., Franken J., Mol J.N.M., van Tunen A.J.;
RT   "Differential expression of two MADS box genes in wild-type and
RT   mutant petunia flowers.";
RL   Plant Cell 4:983-993(1992).
RN   [2]
RP   REVISIONS TO C-TERMINUS.
RA   Angenent G.C.;
RL   Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M91666; AAA86854.1; -.
DR   PIR; JQ1690; JQ1690.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03093; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       98    170       K-BOX.
SQ   SEQUENCE   241 AA;  27562 MW;  B575BE9468D310BB CRC64;

  Query Match             38.8%;  Score 472;  DB 1;  Length 241;
  Best Local Similarity   42.6%;  Pred. No. 7.8e-25;
  Matches  107;  Conservative   47;  Mismatches   69;  Indels   28;  Gaps    6;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||:|||||||||||||||:|||:||||||:|:||||||||||:||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVL----EHAKLKARVEVLEKNKRNFMGED 116
           | | : |||| :  |   :    ::|  |   :    |: ||||| | |::::|| :|||
Db      61 SSMLKTLERYQKCNYGAPE---TNISTREALEISSQQEYLKLKARYEALQRSQRNLLGED 117

Qy     117 LDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKERE 176
           |  |: |||:||| ||| ::| ||| : | | : :  ||:|: || : | :| ::: |
Db     118 LGPLNSKELESLERQLDMSLKQIRSTRTQLMLDQLQDLQRKEHALNEANRTLKQRLMEGS 177

Qy     177 KKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVE--------RVGGEN-----GGASSL 223
               | :       |:  |   :    :  |||           ::| :|     |||
Db     178 TLNLQWQ------QNAQDVGYGRQATQTQGDGFFHPLECEPTLQIGYQNDPITVGGAG-- 229

Qy     224 TEPNSLLPAWM 234
             |: :  |:
Db     230 PSVNNYMAGWL 240


RESULT 11
AGL4_ARATH
ID   AGL4_ARATH     STANDARD;      PRT;   250 AA.
AC   P29384;
DT   01-DEC-1992 (Rel. 24, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL4.
GN   AGL4 OR AT3G02310 OR F14P3.4.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91160981; PubMed=1672119;
RA   Ma H., Yanofsky M.F., Meyerowitz E.M.;
RT   "AGL1-AGL6, an Arabidopsis gene family with similarity to floral
RT   homeotic and transcription factor genes.";
RL   Genes Dev. 5:484-495(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016720; PubMed=11130713;
RA   Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA   Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA   Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA   De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA   Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA   Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA   Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA   Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA   Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA   Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA   Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA   Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA   Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA   de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA   Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA   Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA   Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA   Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA   Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA   Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA   Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA   Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA   Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT   "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:820-822(2000).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY DURING FLOWER DEVELOPMENT.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; M55552; AAA32734.1; -.
DR   EMBL; AC009755; AAF02125.1; -.
DR   PIR; D39534; D39534.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03027; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       97    169       K-BOX.
SQ   SEQUENCE   250 AA;  28577 MW;  53136FE263389943 CRC64;

  Query Match             38.1%;  Score 463.5;  DB 1;  Length 250;
  Best Local Similarity   41.1%;  Pred. No. 3e-24;
  Matches  109;  Conservative   40;  Mismatches   71;  Indels   45;  Gaps    5;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||:|||||||||||||||:|||:||||||:|:||||||||:|||||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVSLIVFSNRGKLYEFCST 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           | | : |||| :  |   ::  :   : ||   |: ||| | | |:: :|| :|||| |
Db      61 SNMLKTLERYQKCSYGSIEVNNKPAKELENSYREYLKLKGRYENLQRQQRNLLGEDLGPL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKE-----R 175
           : |||: || ||| ::| :|  | | | : :| || |:  | | | :|  |:::
Db     121 NSKELEQLERQLDGSLKQVRCIKTQYMLDQLSDLQGKEHILLDANRALSMKLEDMIGVRH 180

Qy     176 EKKTGQQEG------------------QLVQC------SNSSSVLLPQYCVTSSRDGFV 210
             |  ||                    | ::|      |  |   |  ||
Db     181 HHIGGGWEGGDQQNIAYGHPQAHSQGLYQSLECDPTLQIGYSHPVCSEQMAVT------- 233

Qy     211 ERVGGENGGASSLTEPNSLLPAWML 235
           | |::       : |  :| |||
Db     234 --VQGQS------QQGNGYIPGWML 250


RESULT 12
AGL9_ARATH
ID   AGL9_ARATH     STANDARD;      PRT;   251 AA.
AC   O22456;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL9.
GN   AGL9 OR AT1G24260 OR F3I6.19.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RA   Mandel M.A., Yanofsky M.F.;
RT   "The Arabidopsis AGL9 MADS box gene is expressed in young flower
RT   primordia.";
RL   Sex. Plant Reprod. 11:22-28(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016719; PubMed=11130712;
RA   Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA   White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
RA   Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
RA   Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA   Dunn P., Etgu P., Feldblyum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA   Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizar L.,
RA   Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA   Kim C.J., Koo H.L., Kremenetskaia I., Kurtz D.B., Kwan A., Lam B.,
RA   Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA   Lin X., Liu S.X., Liu Z.A., Luros J.S., Maiti R., Marziali A.,
RA   Militscher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA   Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA   Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA   Sun H., Tallon L.J., Tambunga G., Toriumi M.J., Town C.D.,
RA   Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
RA   Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT   "Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:816-820(2000).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR ACTIVE IN INFLORESCENCE
CC       DEVELOPMENT AND FLORAL ORGANOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY DURING FLOWER DEVELOPMENT
CC       WITHIN PETALS, STAMENS, AND CARPELS.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF015552; AAB67832.1; -.
DR   EMBL; AC002396; AAC00586.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03032; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN      100    172       K-BOX.
SQ   SEQUENCE   251 AA;  29066 MW;  0057CABD3F1AFC40 CRC64;

  Query Match             37.8%;  Score 461;  DB 1;  Length 251;
  Best Local Similarity   41.0%;  Pred. No. 4.4e-24;
  Matches  107;  Conservative   47;  Mismatches   75;  Indels   32;  Gaps    6;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||:||||||||||||:|||:||||||:|:||||||||:|||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy      61 SCMERILERYDRYLYS-------DKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFM 113
           | | | |||| :  |         :: :  ::|  :    |: ||| | : |:: :|| :
Db      61 SSMLRTLERYQKCNYGAPEPNVPSREALAVELSSQQ----EYLKLKERYDALQRTQRNLL 116

Qy     114 GEDLDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSL----- 168
           ||||  || |||:||| |||:::| ||: : | | : :: || |:: | : | :|
Db     117 GEDLGPLSTKELESLERQLDSSLKQIRALRTQFMLDQLNDLQSKERMLTETNKTLRLRLA 176

Qy     169 --------LKKIKEREKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGA 220
                   |   :|     |:   | |  :| :   | |    : |:  :  |  |
Db     177 DGYQMPLQLNPNQEEVDHYGRHHHQ--QQQHSQAFFQPLECEPILQIGYQGQQDGMGAGP 234

Qy     221 SSLTEPNSLLPAWMLRPTTTN 241
           |     |: :  |:  |  ||
Db     235 S----VNNYMLGWL--PYDTN 249
```

```
RESULT 13
AGL9_SINAL
ID   AGL9_SINAL     STANDARD;      PRT;   254 AA.
AC   O04067;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL9 homolog (MADS D).
GN   AGL9.
OS   Sinapis alba (White mustard) (Brassica hirta).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Sinapis.
OX   NCBI_TaxID=3728;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Flower;
RX   MEDLINE=97390682; PubMed=9247539;
RA   Bonhomme F., Sommer H., Bernier G., Jacqmard A.;
RT   "Characterization of SaMADS D from Sinapis alba suggests a dual
RT   function of the gene: in inflorescence development and floral
RT   organogenesis.";
RL   Plant Mol. Biol. 34:573-573(1997).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR ACTIVE IN INFLORESCENCE
CC       DEVELOPMENT AND FLORAL ORGANOGENESIS.
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; Y08626; CAA69916.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03170; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN      100    172       K-BOX.
SQ   SEQUENCE   254 AA;  29603 MW;  EDD65155060BFD46 CRC64;

  Query Match             36.9%;  Score 449.5;  DB 1;  Length 254;
  Best Local Similarity   39.3%;  Pred. No. 2.6e-23;
  Matches  103;  Conservative   53;  Mismatches   75;  Indels   31;  Gaps    6;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||:|||||||||||||||:|||:||||||:|:|||||||||||:||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy      61 SCMERILERYDRYLYS-------DKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFM 113
           | | | |||| :  |         :: :  ::|  :    |: ||| | : |:: :|| :
Db      61 SSMIRTLERYQKCNYGPPEPNVPSREALAVELSSQQ----EYLKLKERYDALQRTQRNLL 116

Qy     114 GEDLDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKI- 172
           ||||  || |||: || |||:::| ||: : | | : :: || |:: | : | :|  ::
Db     117 GEDLGPLSTKELELLERQLDSSLKQIRALRTQFMLDQLNDLQSKERMLNETNKTLRLRLA 176

Qy     173 ------------KEREKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGG 219
                       ::     |: : |  | |:  :   |  |    : |:    | :: |

Db     177 DGYQMPLQLNPNQEDHHVDYGRHDQQQQQNSH-HAFFQPLECEPILQMGY---QGQQDHG 232

Qy     220 ASSLTEPNSLLPAWMLRPTTTN 241
             :    |: :  |:  |  ||
Db     233 MEAGPSENNYMLGWL--PYDTN 252


RESULT 14
AGL9_LYCES
ID   AGL9_LYCES     STANDARD;      PRT;   224 AA.
AC   Q42464;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL9 homolog (TM5).
GN   TDR5.
OS   Lycopersicon esculentum (Tomato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4081;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. VFNT Cherry; TISSUE=Flower;
RX   MEDLINE=93251098; PubMed=1688249;
RA   Pnueli L., Abu-Abeid M., Zamir D., Nacken W., Schwarz-Sommer Z.,
RA   Lifschitz E.;
RT   "The MADS box gene family in tomato: temporal expression during
RT   floral development, conserved secondary structures and homology with
RT   homeotic genes from Antirrhinum and Arabidopsis.";
RL   Plant J. 1:255-266(1991).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR ACTIVE IN INFLORESCENCE
CC       DEVELOPMENT AND FLORAL ORGANOGENESIS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- TISSUE SPECIFICITY: FLOWER-SPECIFIC.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X60758; CAA43170.1; -.
DR   EMBL; X60480; CAA43010.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03187; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       98    170       K-BOX.
SQ   SEQUENCE   224 AA;  25999 MW;  51D10D30867D06F7 CRC64;

  Query Match             36.9%;  Score 449;  DB 1;  Length 224;
  Best Local Similarity   48.7%;  Pred. No. 2.4e-23;
  Matches   96;  Conservative   41;  Mismatches   40;  Indels   20;  Gaps    5;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||:||||| |||||||||:|||:||||||:|:|||||||||||:||::|||:|: :
Db       1 MGRGRVELKRIEGKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy      61 SCMERILERYDRYLY-------SDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFM 113
```

```
          | | : |||| :  |       | :: :  ::|  :    |: ||| | | |::::|| :
Db     61 SSMLKTLERYQKCNYGAPEPNISTREAL--EISSQQ----EYLKLKGRYEALQRSQRNLL 114

Qy    114 GEDLDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIK 173
          ||||  |: |||:||| ||| ::| ||| : | | : ::  |:|: || : | :| :::
Db    115 GEDLGPLNSKELESLERQLDMSLKQIRSTRTQLMLDQLTDYQRKEHALNEANRTLKQRLM 174

Qy    174 EREKKTGQQEGQLVQCS 190
          |     | |    :|||
Db    175 E-----GSQLN--LQCS 184


RESULT 15
AGL9_ARADE
ID   AGL9_ARADE     STANDARD;      PRT;   250 AA.
AC   Q38694;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-JUN-2002 (Rel. 41, Last annotation update)
DE   Agamous-like MADS box protein AGL9 homolog (OM1).
OS   Aranda deborah (Orchid).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
OC   Epidendroideae; higher Epidendroideae; Vandeae; Aeridinae;
OC   Arachnis x Vanda.
OX   NCBI_TaxID=29714;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94072738; PubMed=8251643;
RA   Lu Z.X., Wu M., Loh C.S., Yeong C.Y., Goh C.J.;
RT   "Nucleotide sequence of a flower-specific MADS box cDNA clone from
RT   orchid.";
RL   Plant Mol. Biol. 23:901-904(1993).
CC   -!- FUNCTION: PROBABLE TRANSCRIPTION FACTOR ACTIVE IN INFLORESCENCE
CC       DEVELOPMENT AND FLORAL ORGANOGENESIS (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN PETALS AND WEAKLY IN SEPALS BUT
CC       NOT IN THE COLUMN (GYNOSTEMIUM).
CC   -!- DEVELOPMENTAL STAGE: EXPRESSED IN MATURE FLOWERS AND NOT IN YOUNG
CC       DEVELOPING INFLORESCENCES OR YOUNG FLORAL BUDS.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: CONTAINS 1 K-BOX DIMERIZATION DOMAIN.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X69107; CAA48859.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03114; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF00319; SRF-TF; 1.
DR   Pfam; PF01486; K-box; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Transcription regulation; DNA-binding; Nuclear protein.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       96    168       K-BOX.
SQ   SEQUENCE   250 AA;  28770 MW;  01EF94DADC499C41 CRC64;

  Query Match             36.6%;  Score 446;  DB 1;  Length 250;
  Best Local Similarity   38.9%;  Pred. No. 4.3e-23;
  Matches   96;  Conservative   56;  Mismatches   77;  Indels   18;  Gaps    4;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||||:|| |||||||||||:|||  |||||:|:|||||||||||:||::|||:|: :
Db      1 MGRGRVELKMIENKINRQVTFAKRRKRLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy     61 SCMERILERYDRYLYS--DKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLD 118
          : | : ||:| :  :   :  :: |:   |:    |: ||| ||| |::::|| :||||
Db     61 TSMLKTLEKYQKCNFGSPESTIISRETQSSQQ---EYLKLKNRVEALQRSQRNLLGEDLG 117

Qy    119 SLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKK 178
           |  |||: || |||:::: ||| : | | : :: ||:::: | : | :| :: :|  :
Db    118 PLGSKELEQLERQLDSSLRQIRSTRTQFMLDQLADLQRREQMLCEANKTLKRRFEESSQA 177

Qy    179 TGQQ----------EGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPN 227
            ||            |:     :  :   |  |  : : |:   :      ||::   |
Db    178 NQQQVWDPSNTHAVGYGRQPAQHHGEAFYHPLECEPTLQIGYHSDITMATATASTVN--N 235

Qy    228 SLLPAWM 234
          : | |:
Db    236 YMPPGWL 242


Search completed: January 27, 2003, 12:38:22
Job time : 13 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 27, 2003, 12:37:27 ; Search time 19 Seconds
(without alignments)
1224.449 Million cell updates/sec

Title: US-09-981-087A-2
Perfect score: 1218
Sequence: 1 MGRGRVQLKRIENKINRQVT.....LTEPNSLLPAWMLRPTTTNE 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1218	100.0	242	2	S71208	MADS box protein A
2	743	61.0	250	2	T07100	MADS box protein h
3	731	60.0	250	2	T07902	MADS box protein -
4	709.5	58.3	227	2	S23730	MADS box protein T
5	683	56.1	248	2	S20886	MADS box protein s
6	662.5	54.4	256	2	S27109	MADS box protein A
7	656	53.9	254	2	S52236	MADS box protein a
8	621.5	51.0	251	2	T14456	MADS box protein h
9	610.5	50.1	273	2	T03410	MADS box protein -
10	575.5	47.2	228	2	T14737	MADS box protein -
11	548	45.0	150	2	T14457	MADS box protein h
12	509	41.8	233	2	T10714	MADS-box protein C
13	500	41.1	245	2	T09569	MADS box protein M
14	497.5	40.8	261	2	S51935	probable MADS-box
15	497.5	40.8	261	2	T09603	MADS-box protein 3
16	497	40.8	242	2	T10486	MADS box protein -
17	497	40.8	242	2	T09571	MADS box protein M
18	491.5	40.4	246	2	T17023	MADS box protein 1
19	491	40.3	257	2	S53306	MADS box protein M
20	484.5	39.8	248	2	B39534	MADS box protein A
21	476	39.1	249	2	T04335	MADS box protein -
22	473	38.8	248	2	T04170	MADS box protein -
23	470.5	38.6	225	2	T04168	MADS box protein -
24	470	38.6	249	2	T04307	M79 protein - rice
25	467	38.3	231	2	T14801	MADS box protein M
26	466	38.3	247	2	S78015	MADS box protein D
27	466	38.3	259	2	T04169	MADS box protein -
28	464.5	38.1	262	2	T51409	MADS box protein A
29	463.5	38.1	242	2	S71757	MADS box protein D
30	463.5	38.1	250	2	D39534	MADS box protein A
31	462	37.9	243	2	S71756	MADS box protein D
32	461	37.8	251	2	T00656	MADS box protein A
33	460	37.8	224	2	JQ1690	MADS box protein f
34	456	37.4	250	2	T04167	MADS box protein -
35	455.5	37.4	255	2	T03408	MADS box protein -
36	454	37.3	255	2	T03398	MADS box protein -
37	453.5	37.2	247	2	T06543	MADS box protein -
38	449.5	36.9	254	2	T10467	MADS box protein D
39	449	36.9	224	2	S23728	MADS box protein T
40	446	36.6	250	2	S40405	MADS box protein o
41	442	36.3	252	2	F39534	floral homeotic pr
42	432.5	35.5	258	2	S57793	MADS box protein A
43	427	35.1	244	2	T47904	MADS-box protein A
44	420	34.5	219	2	S46526	MADS box protein m
45	416.5	34.2	248	2	T03592	floral homeotic pr

ALIGNMENTS

RESULT 1
S71208
MADS box protein AGL8 - Arabidopsis thaliana
N;Alternate names: agamous-like protein 8
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 28-Oct-1996 #sequence_revision 27-Feb-1997 #text_change 24-Sep-1999
C;Accession: S71208
R;Mandel, M.A.; Yanofsky, M.F.
submitted to the EMBL Data Library, August 1995
A;Description: The Arabidopsis AGL8 MADS box gene is expressed in inflorescence meris
A;Reference number: S71208
A;Accession: S71208
A;Molecule type: mRNA
A;Residues: 1-242 <MAN>
A;Cross-references: EMBL:U33473; NID:g1004364; PIDN:AAA97403.1; PID:g1004365
C;Genetics:
A;Gene: AGL8
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain ho
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 100.0%; Score 1218; DB 2; Length 242;
Best Local Similarity 100.0%; Pred. No. 5.2e-73;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180

Qy     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPAWMLRPTTT 240

Qy     241 NE 242
           ||
Db     241 NE 242
```

RESULT 2
T07100
MADS box protein homolog POTM1-1 - potato
C;Species: Solanum tuberosum (potato)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 11-May-2000
C;Accession: T07100

```
R;Kang, S.G.; Hannapel, D.J.
Gene 166, 329-330, 1995
A;Title: Nucleotide sequences of novel potato (Solanum tuberosum L.) MADS-box cDNAs and
A;Reference number: Z15919; MUID:96125213; PMID:8543185
A;Accession: T07100
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-250 <KAN>
A;Cross-references: EMBL:U23757; NID:g758564; PIDN:AAA92839.1; PID:g758565
A;Experimental source: cv. Superior; vegetative tissue; tuber induced 4-day axillary bud
C;Genetics:
A;Gene: POTM1-1
C;Function:
A;Description: may be involved in vegetative organ development of potato
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription factor; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             61.0%;  Score 743;  DB 2;  Length 250;
  Best Local Similarity   61.8%;  Pred. No. 7.5e-42;
  Matches  155;  Conservative   36;  Mismatches   40;  Indels   20;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||||||||||||||||||||| |||||:|||||||: |
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYAND 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           |||||:||||:|| ::::|||  | :  :| |||||||||:|||::|:::::||||:|| 
Db      61 SCMERLLERYERYSFAERQLVPTDHTSPGSWTLEHAKLKARLEVLQRNQKHYVGEDLESL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ::|||||:|||||||:|:| ||||||| | |||| |||:|:|||: || | ||:|||||:
Db     121 NMKELQNLEHQLDSALKHIRSRKNQLMHESISVLQKQDRALQEQNNQLSKKVKEREKEVA 180

Qy     181 QQE--GQLVQCSNSSSVLLPQ----------YCVTSSRDGFVERVGGENGGASSLTE--- 225
           ||    |     |||: :|||          |  |:  |      |   || ||  :
Db     181 QQNQWDQQNHEINSSTFVLPQQLDSPHLGEAYQNTNVVDN-----GEVEGGNSSQQQGAA 235

Qy     226 PNSLLPAWMLR 236
           |:::| ||||
Db     236 NNTVMPQWMLR 246


RESULT 3
T07902
MADS box protein - Commerson's wild potato
C;Species: Solanum commersonii (Commerson's wild potato)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 08-Oct-1999
C;Accession: T07902
R;Seppanen, M.M.
submitted to the EMBL Data Library, May 1997
A;Description: Nucleotide sequence of novel wild potato (Solanum commersonii) MADS-box c
A;Reference number: Z16201
A;Accession: T07902
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-250 <SEP>
A;Cross-references: EMBL:AF002666; NID:g2290777; PIDN:AAB65161.1; PID:g2290778
C;Genetics:
A;Gene: SCM1
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             60.0%;  Score 731;  DB 2;  Length 250;
  Best Local Similarity   62.2%;  Pred. No. 4.6e-41;
  Matches  153;  Conservative   39;  Mismatches   44;  Indels   10;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||||||||||||||||||||| |||||:|||||||: |
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           |||||:||||:|| ::::|||  | :  :| |||||||||:|||::|:::::||||:|| 
Db      61 SCMERLLERYERYSFAERQLVPTDHTSPGSWTLEHAKLKARLEVLQRNQKHYVGEDLESL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ::|||||:|||||||:|:| ||||||| | |||| |||:|:|||: || | ||:|||||:
Db     121 NMKELQNLEHQLASALKHIRSRKNQLMHESISVLQKQDRALQEQNNQLSKKVKEREKEVE 180

Qy     181 QQE--GQLVQCSNSSSVLLPQ-----YCVTSSRDGFVERVGGENGGASSLTE---PNSLL 230
           ||    |     |||: :|||     :   :|::  |   |   || ||  :    |:::
Db     181 QQNQWDQQNHEINSSTFVLPQQLDSPHLGEASQNTNVVDNGEVEGGNSSQXQGAANNTVM 240

Qy     231 PAWMLR 236
           | ||:|
Db     241 PQWMVR 246


RESULT 4
S23730
MADS box protein TDR4 - tomato
N;Alternate names: floral homeotic protein TM4
C;Species: Lycopersicon esculentum (tomato)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 24-Sep-1999
C;Accession: S23730
R;Pnueli, L.; Abu-Abeid, M.; Zamir, D.; Nacken, W.; Schwarz-Sommer, Z.; Lifschitz, E.
Plant J. 1, 255-266, 1991
A;Title: The MADS box gene family in tomato: temporal expression during floral develo
A;Reference number: S23728; MUID:93251098; PMID:1688249
A;Accession: S23730
A;Molecule type: mRNA
A;Residues: 1-227 <PNU>
A;Cross-references: EMBL:X60757; NID:g19381; PIDN:CAA43169.1; PID:g19382
C;Genetics:
A;Gene: TDR4
A;Map position: 6
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain ho
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             58.3%;  Score 709.5;  DB 2;  Length 227;
  Best Local Similarity   72.3%;  Pred. No. 1.1e-39;
  Matches  141;  Conservative   25;  Mismatches   24;  Indels    5;  Gaps    1;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||||||||||||||||||||| |||||:|||||||: |
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYAND 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| :::|||| | :  :| ||| |||||:|||::|:::::||||:||
Db      61 SCMERILERYERYSFAEKQLVPTDHTSPVSWTLEHRKLKARLEVLQRNQKHYVGEDLESL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |:||||:|||||:|:| |||||||| | |||| |||||:|||: || | ||:|||||
Db     121 SMKELQNLEHQLDSALKHIRSRKNQLMHESISVLQKKDRALQEQNNQLSKKVKEREKSAQ 180

Qy     181 QQEGQLVQCSNSSSV 195
           |  |      ||||:
Db     181 QISG-----INSSSL 190


RESULT 5
S20886
MADS box protein squa - garden snapdragon
N;Alternate names: transcription factor squa
C;Species: Antirrhinum majus (garden snapdragon)
C;Date: 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change 24-Sep-1999
C;Accession: S20886
R;Huijser, P.; Klein, J.; Loennig, W.E.; Meijer, H.; Saedler, H.; Sommer, H.
EMBO J. 11, 1239-1249, 1992
A;Title: Bracteomania, an inflorescence anomaly, is caused by the loss of function of
A;Reference number: S20886; MUID:92224862; PMID:1563342
A;Accession: S20886
A;Molecule type: DNA
A;Residues: 1-248 <HUI>
```

```
A;Cross-references: EMBL:X63701; NID:g16051; PIDN:CAA45228.1; PID:g16052
C;Genetics:
A;Gene: sqa
A;Introns: 62/2; 88/3; 110/2; 143/3; 157/3; 171/3; 215/2
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             56.1%;  Score 683;  DB 2;  Length 248;
  Best Local Similarity   70.3%;  Pred. No. 6.4e-38;
  Matches  128;  Conservative   32;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||:||||||||||||||||||| |||||||:|||||||||||||:||||||||||
Db      1 MGRGKVQLKRIENKINRQVTFSKRRGGLLKKAHELSVLCDAEVALIVFSNKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          |||:||||:|:|| ::::||| :     || ||::|||||:|:|::| |::|||||||:
Db     61 SCMDRILEKYERYSFAERQLVSNEPQSPANWTLEYSKLKARIELLQRNHRHYMGEDLDSM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||:|||| ||| |:|:||:|||| :::||| || |:||:|: |  | |||||:||:
Db    121 SLKEIQSLEQQLDTALKNIRTRKNQLLYDSISELQHKEKAIQEQNTMLAKKIKEKEKEIA 180

Qy    181 QQ 182
          ||
Db    181 QQ 182


RESULT 6
S27109
MADS box protein AP1 - Arabidopsis thaliana
N;Alternate names: floral homeotic protein AP1
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Sep-1999
C;Accession: S27109; S36597
R;Mandel, M.A.; Gustafson-Brown, C.; Savidge, B.; Yanofsky, M.F.
Nature 360, 273-277, 1992
A;Title: Molecular characterization of the Arabidopsis floral homeotic gene APETALA1.
A;Reference number: S27109; MUID:93063372; PMID:1359429
A;Accession: S27109
A;Molecule type: mRNA
A;Residues: 1-256 <MAN>
A;Cross-references: EMBL:Z16421; NID:g16161; PIDN:CAA78909.1; PID:g16162
A;Note: 27-Asp was also found
R;Yanofsky, M.F.
submitted to the EMBL Data Library, October 1992
A;Reference number: S36597
A;Accession: S36597
A;Molecule type: mRNA
A;Residues: 1-235,237-256 <YAN>
A;Cross-references: EMBL:Z16421; NID:g16161; PIDN:CAA78909.1; PID:g16162
C;Genetics:
A;Map position: 1
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             54.4%;  Score 662.5;  DB 2;  Length 256;
  Best Local Similarity   65.8%;  Pred. No. 1.5e-36;
  Matches  133;  Conservative   33;  Mismatches   33;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          |||||||||||||||||||||||:|||||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|::::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREK-KT 179
          | ||||:|| ||| |:| ||:|||| |:|||: ||||:||:|: |: | |:||||||
Db    121 SPKELQNLEQQLDTALKHIRTRKNQLMYESINELQKKEKAIQEQNSMLSKQIKEREKILR 180

Qy    180 GQQE--GQLVQCSNSSSVLLPQ 199
          |||   |  |  |    | ||
Db    181 AQQEQWDQQNQGHNMPPPLPPQ 202


RESULT 7
S52236
MADS box protein ap1 - white mustard
N;Alternate names: floral homeotic protein ap1
C;Species: Sinapis alba (white mustard)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 24-Sep-1999
C;Accession: S52236
R;Menzel, G.; Apel, K.; Melzer, S.
submitted to the EMBL Data Library, September 1994
A;Description: Isolation and sequencing of the mustard (Sinapis alba L.) ap1 cDNA.
A;Reference number: S52236
A;Accession: S52236
A;Molecule type: mRNA
A;Residues: 1-254 <MEN>
A;Cross-references: EMBL:X81480; NID:g609252; PIDN:CAA57233.1; PID:g609253
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain ho
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             53.9%;  Score 656;  DB 2;  Length 254;
  Best Local Similarity   64.1%;  Pred. No. 3.9e-36;
  Matches  127;  Conservative   38;  Mismatches   33;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          |||||||||||||||||||||||||:||||||||||||||||||:||| ||||||||||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALVVFSHKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::||:  :   : || :|: :|||::|:||:|:|::::|||| ::
Db     61 SCMEKILERYERYSYAERQLIAPESDVNTNWSMEYNRLKAKIELLERNQRHYLGEDLQAM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          | ||||:|| ||| |:| ||||||| | :||: ||:|:||:|: |: | |:||||||
Db    121 SSKELQNLEQQLDTALKHIRSRKNQLMHDSINELQRKEKAIQEQNSMLSKQIKEREKILR 180

Qy    181 QQEGQLVQCSNSSSVLLP 198
          |: |  | ::  ::  |
Db    181 AQQEQWDQQNHGHNMPPP 198


RESULT 8
T14456
MADS box protein homolog CAL - wild cabbage
C;Species: Brassica oleracea (wild cabbage)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14456
R;Kempin, S.A.; Savidge, B.; Yanofsky, M.F.
Science 267, 522-525, 1995
A;Title: Molecular basis of the cauliflower phenotype in Arabidopsis.
A;Reference number: Z18101; MUID:95125463; PMID:7824951
A;Accession: T14456
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-251 <KEM>
A;Cross-references: EMBL:L36926; NID:g642590; PIDN:AAA64790.1; PID:g642591
A;Experimental source: floral meristem
C;Genetics:
A;Gene: CAL
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain ho
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             51.0%;  Score 621.5;  DB 2;  Length 251;
  Best Local Similarity   60.2%;  Pred. No. 7e-34;
  Matches  121;  Conservative   42;  Mismatches   35;  Indels    3;  Gaps    2;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
```

```
          |||||||::|||||||||||||||||||:||||||||||:||||||:||||| ||||||||::
Db        1 MGRGRVEMKRIENKINRQVTFSKRRAGLLKKAHEISILCDAEVSLIVFSHKGKLFEYSSE 60

Qy       61 SCMERILERYDRYLYSDKQLVGRD--VSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLD 118
            ||||::|| |:|| |::|||   |  |:   || :|:::|||::|:||:|:|:::||||:
Db       61 SCMEKVLEHYERYSYAEKQLKVPDSHVNAQTNWSVEYSRLKAKIELLERNQRHYLGEDLE 120

Qy      119 SLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKK 178
            |:|:||||:|| ||| ::| ||||||| | ||:: ||:|:|: : |: | |:|:|||
Db      121 SISIKELQNLEQQLDTSLKHIRSRKNQLMHESLNHLQRKEKEILEENSMLAKQIRERESI 180

Qy      179 TGQQEGQLVQCSNSSSVLLPQ 199
               : |  |  | |  : ||
Db      181 LRTHQNQSEQ-QNRSHHVAPQ 200


RESULT 9
T03410
MADS box protein - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 24-Sep-1999
C;Accession: T03410
R;Mena, M.; Mandel, M.A.; Lerner, D.R.; Yanofsky, M.F.; Schmidt, R.J.
Plant J. 8, 845-854, 1995
A;Title: A characterization of the MADS-box gene family in maize.
A;Reference number: Z14928; MUID:96132144; PMID:8580958
A;Accession: T03410
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-273 <MEN>
A;Cross-references: EMBL:L46400; NID:g939784; PIDN:AAB00081.1; PID:g939785
C;Genetics:
A;Note: ZAP1
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             50.1%;  Score 610.5;  DB 2;  Length 273;
  Best Local Similarity   49.1%;  Pred. No. 4.1e-33;
  Matches  133;  Conservative   40;  Mismatches   59;  Indels   39;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
            ||||:|||||||||||||||||:|||||||||||||||||:|||| ||||:||:||
Db       1 MGRGKVQLKRIENKINRQVTFSKRRNGLLKKAHEISVLCDAEVAVIVFSPKGKLYEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
            | |::|||||:|| |::| |:  :     ||  |: ||||::| ::  :: |||||:||
Db      61 SRMDKILERYERYSYAEKALISAESESEGNWCHEYRKLKAKIETIQKCHKHLMGEDLESL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
            : |||| || |||:::| |||||:  | |||| ||||:::||: | :| |:: ||:|
Db     121 NPKELQQLEQQLDSSLKHIRSRKSHLMAESISELQKKERSLQEENKALQKELAERQKAVA 180

Qy     181 ---QQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERV----------------------GGE 216
               ||: | ||    :     |   :||   |: |                       | |
Db     181 SRQQQQQQQVQWDQQTHA---QAQTSSSSSSFMMRQDQQGLPPPHNICFPPLTMGDRGEE 237

Qy     217 NGGASSLTEPNSL------------LPAWML 235
           |::  :  |              || |||
Db     238 LAAAAAAQQQQPLPGQAQPQLRIAGLPPWML 268


RESULT 10
T14737
MADS box protein - sorghum (fragment)
C;Species: Sorghum bicolor (sorghum)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14737
R;Greco, R.; Stagi, L.; Colombo, L.; Angenent, G.C.; Sari-Gorla, M.; Pe, M.E.
Mol. Gen. Genet. 253, 615-623, 1997
A;Title: MADS box genes expressed in developing inflorescences of rice and sorghum.
A;Reference number: Z15292; MUID:97218034; PMID:9065695
A;Accession: T14737
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-228 <GRE>
A;Cross-references: EMBL:U32110; NID:g1905933; PIDN:AAB50181.1; PID:g1905934
C;Genetics:
A;Gene: MADS2
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain ho
C;Keywords: DNA binding; nucleus; transcription regulation

  Query Match             47.2%;  Score 575.5;  DB 2;  Length 228;
  Best Local Similarity   51.9%;  Pred. No. 6.5e-31;
  Matches  120;  Conservative   40;  Mismatches   58;  Indels   13;  Gaps    3;

Qy       5 RVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTDSCME 64
           :|||||||||||||||||||:||||||||||||||||:|||| ||||:||:||| |:
Db       1 KVQLKRIENKINRQVTFSKRRNGLLKKAHEISVLCDAEVAVIVFSPKGKLYEYATDSRMD 60

Qy      65 RILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSLSLKE 124
           :|||||:|| |::| |:  :     ||  |: ||||::| ::|  :: |||||:||: ||
Db      61 KILERYERYSYAEKALISAESESEGNWCHEYRKLKAKIETIQKCHKHLMGEDLESLNPKE 120

Qy     125 LQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG---Q 181
           || || ||:::| |||||:  | |||| ||||:::||: | :| |:: ||:|      |
Db     121 LQQLEQQLESSLKHIRSRKSHLMAESISELQKKERSLQEENKALQKELAERQKAAASRQQ 180

Qy     182 QEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSLLPA 232
           |:| :          |   :       :  :  | | | |      :|||
Db     181 QQGAVGPADTDPG---PDKLI-------IVLLHDEAGSAGSAASTKHMLPA 221


RESULT 11
T14457
MADS box protein homolog CAL - broccoli
C;Species: Brassica oleracea var. botrytis (broccoli)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14457
R;Kempin, S.A.; Savidge, B.; Yanofsky, M.F.
Science 267, 522-525, 1995
A;Title: Molecular basis of the cauliflower phenotype in Arabidopsis.
A;Reference number: Z18101; MUID:95125463; PMID:7824951
A;Accession: T14457
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-150 <KEM>
A;Cross-references: EMBL:L36927; NID:g642592; PIDN:AAA64791.1; PID:g642593
A;Experimental source: curd (inflorescence meristem)
C;Genetics:
A;Gene: CAL
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain ho
C;Keywords: DNA binding; nucleus; transcription regulation

  Query Match             45.0%;  Score 548;  DB 2;  Length 150;
  Best Local Similarity   68.5%;  Pred. No. 2.5e-29;
  Matches  102;  Conservative   31;  Mismatches   14;  Indels    2;  Gaps    1;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||::|||||||||||||||||:||||||||:||||||:||||| |||||||::
Db       1 MGRGRVEMKRIENKINRQVTFSKRRAGLLKKAHEISILCDAEVSLIVFSHKGKLFEYSSE 60

Qy      61 SCMERILERYDRYLYSDKQLVGRD--VSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLD 118
           ||||::|||||:|| |::|||   |  |:   || :|:::|||::|: |:|:|:::|||||:
Db      61 SCMEKVLERYERYSYAEKQLKAPDSHVNAQTNWSMEYSRLKAKIELWERNQRHYLGEDLE 120

Qy     119 SLSLKELQSLEHQLDAAIKSIRSRKNQAM 147
           |:|:||||:|| ||| ::| ||||||| |
Db     121 SISIKELQNLEQQLDTSLKHIRSRKNQLM 149


RESULT 12
T10714
MADS-box protein CMB1 - clove pink
C;Species: Dianthus caryophyllus (clove pink)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
C;Accession: T10714
```

```
R;Baudinette, S.C.; Savin, K.W.
submitted to the EMBL Data Library, March 1995
A;Description: Carnation MADS box genes.
A;Reference number: Z17094
A;Accession: T10714
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-233 <BAU>
A;Cross-references: EMBL:L40404; NID:g695316; PID:g695317
A;Experimental source: cv. Scania; petals
C;Genetics:
A;Gene: CMB1
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; transcription factor; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             41.8%;  Score 509;  DB 2;  Length 233;
  Best Local Similarity   44.3%;  Pred. No. 1.5e-26;
  Matches  109;  Conservative   45;  Mismatches   68;  Indels   24;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||:|||||||||||||||:|||:||||||:|:|||||||||||||::|||:|: :
Db       1 MGRGRVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIVFSNRGKLYEFCST 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||| : |||| |  |    :   :   ::|:   |: ||||:|:||::: || :|||  |
Db      61 SCMNKTLERYQRCSYGSLE-TSQPSKETESSYQEYLKLKAKVDVLQRSHRNLLGEDLGEL 119

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKE------ 174
           | |||: |||||| ::: ||| | | | : :: ||||:: | : | :|  |::|
Db     120 STKELEQLEHQLDKSLRQIRSIKTQHMLDQLADLQKKEEMLFESNRALKTKLEESCASFR 179

Qy     175 -----REKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPNSL 229
                |:   |  |   : |:|:  :            |: |    :    :|    :
Db     180 PNWDVRQPGDGFFEPLPLPCNNNLQI-----------GYNEATQDQMNATTSAQNVHGF 227

Qy     230 LPAWML 235
           |||
Db     228 AQGWML 233


RESULT 13
T09569
MADS box protein MADS1 - Monterey pine
C;Species: Pinus radiata (Monterey pine)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
C;Accession: T09569
R;Mouradov, A.; Glassick, T.; Vivian-Smith, A.; Teasdale, R.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z16739
A;Accession: T09569
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-245 <MOU>
A;Cross-references: EMBL:U42399; NID:g1206002; PID:g1206003
C;Genetics:
A;Gene: MADS1
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             41.1%;  Score 500;  DB 2;  Length 245;
  Best Local Similarity   43.2%;  Pred. No. 6.3e-26;
  Matches  108;  Conservative   53;  Mismatches   69;  Indels   20;  Gaps    5;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||:|:|||||||||||||:|||:||||||:|:|||||||||||:||::|||:|: :
Db       1 MGRGKVELKRIENKINRQVTFAKRRNGLLKKAYELSVLCDAEVALIIFSNRGKLYEFCSS 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           | | : :|:| :  |   :     ::: :|   :: :|||||||:::::|| :||:|  |
Db      61 SSMMKTIEKYQKCSYGSLE-TNCSINEMQNSYQDYLELKARVEVLQRSQRNLLGEELGPL 119

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           : |||: |||||: ::| ||| | | ||: :: || |:: | : |  | ||::|   :
Db     120 NSKELEQLEHQLENSLKQIRSAKTQFMFDQLAHLQHKEQMLVEANRELRKKLEESNTRIP 179

Qy     181 QQEGQLVQCSNSSSV-LLPQYCVTSSRDGFVERVGG--------ENGGASSLT------E 225
            : |   :  |: |   ||    | |:    : :||           |:: |
Db     180 LRLGWEAEDHNNISYRRLP----TQSQGLIFQPLGGYPNMQIGYNPAGSNELNVSPADQH 235

Qy     226 PNSLLPAWML 235
           ||  :| |||
Db     236 PNGFIPGWML 245


RESULT 14
S51935
probable MADS-box protein dal1 - Norway spruce
C;Species: Picea abies (Norway spruce)
C;Date: 14-Jul-1995 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
C;Accession: T14846; S51935
R;Tandre, K.; Albert, V.A.; Sundas, A.; Engstroem, P.
Plant Mol. Biol. 27, 69-78, 1995
A;Title: Conifer homologues to genes that control floral development in angiosperms.
A;Reference number: S51934; MUID:95170009; PMID:7865797
A;Accession: T14846
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-261 <TAN>
A;Cross-references: EMBL:X80902; NID:g695685; PIDN:CAA56864.1; PID:g695686
A;Accession: S51935
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 92-156 <TA2>
C;Genetics:
A;Gene: dal1
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain ho
C;Keywords: DNA binding; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             40.8%;  Score 497.5;  DB 2;  Length 261;
  Best Local Similarity   51.3%;  Pred. No. 9.9e-26;
  Matches  101;  Conservative   45;  Mismatches   42;  Indels    9;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||:|||||||||||||||:||||||:|:||||||||||:|||::|||:|:::
Db       1 MGRGRVQLRRIENKINRQVTFSKRRNGLLKKAYELSVLCDAEVALIIFSTRGKLYEFASS 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           | | : ||||::  |: :   |    :::||  |  ||| :||:|:::::|: :||||  |
Db      61 S-MNKTLERYEKCSYAMQDTTGVSDREAQNWHQEVTKLKGKVELLQRSQRHLLGEDLGPL 119

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKK-- 178
           ::|||| || ||: |:  :|||| | | : |  |:::::: | : | || ||: | | :
Db     120 NVKELQQLERQLEVALAHLRSRKTQVMLDQIEELRQRERLLHEVNKSLQKKLSETEGRDV 179

Qy     179 -TGQQEGQLVQCSNSSS 194
            ||     : | ||:::
Db     180 ITG-----IEQTSNTNT 191


RESULT 15
T09603
MADS-box protein 3 - Monterey pine
C;Species: Pinus radiata (Monterey pine)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jan-2000
C;Accession: T09603
R;Mouradov, A.; Loopstra, C.; Southerton, S.; Glassick, T.; Marshal, H.; Teasdale, R.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z16765
A;Accession: T09603
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-261 <MOU>
A;Cross-references: EMBL:U76726; NID:g2160700; PID:g2160701
```

```
A;Experimental source: tissue-type immature male and female cones
C;Genetics:
A;Gene: MADS3
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             40.8%;  Score 497.5;  DB 2;  Length 261;
  Best Local Similarity   51.3%;  Pred. No. 9.9e-26;
  Matches  101;  Conservative   45;  Mismatches   42;  Indels    9;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||:|||||||||||||||||:||||||:|:||||||||||||:||::|||:|:::
Db       1 MGRGRVQLRRIENKINRQVTFSKRRNGLLKKAYELSVLCDAEVALIIFSTRGKLYEFASS 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           | | : ||||::  |: :   |    :::||  |  ||| :||:|::::|: :||||  |
Db      61 S-MNKTLERYEKCSYAMQDTTGVSDREAQNWHQEVTKLKGKVELLQRSQRHLLGEDLGPL 119

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKK-- 178
           ::|||| || ||: |:  :|||| | | : |  |:::::| : | || ||: | | :
Db     120 NVKELQQLERQLEVALTHLRSRKTQVMLDQIEELRQRERLLHEVNKSLQKKLSETEGRDV 179

Qy     179 -TGQQEGQLVQCSNSSS 194
            ||      : | ||:::
Db     180 ITG-----IEQTSNTNT 191


Search completed: January 27, 2003, 12:39:49
Job time : 20 secs
```

GenCore version 5.1.3

OM protein - protein search, using sw model

Run on: January 27, 2003, 12:34:37 ; Search time 33 Seconds
(without alignments)
1511.012 Million cell updates/sec

Title: US-09-981-087A-2
Perfect score: 1218
Sequence: 1 MGRGRVQLKRIENKINRQVT.....LTEPNSLLPAWMLRPTTTNE 242

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_21:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	794	65.2	244	10	Q39401	Q39401 betula verr
2	772.5	63.4	245	10	Q9XHR7	Q9xhr7 nicotiana s
3	769.5	63.2	255	10	Q9ZRA5	Q9zra5 malus domes
4	761.5	62.5	247	10	Q9SEG7	Q9seg7 capsicum an
5	756.5	62.1	245	10	Q9SBQ0	Q9sbq0 petunia hyb
6	753	61.8	246	10	Q9SBQ1	Q9sbq1 petunia hyb
7	751.5	61.7	245	10	Q9ZTY3	Q9zty3 nicotiana t
8	741	60.8	257	10	Q41356	Q41356 silene lati
9	735.5	60.4	244	10	Q9FUI3	Q9fui3 eucalyptus
10	732.5	60.1	245	10	Q9FUH8	Q9fuh8 eucalyptus
11	703.5	57.8	205	10	Q9FUI2	Q9fui2 eucalyptus
12	695	57.1	252	10	Q941M9	Q941m9 antirrhinum
13	683	56.1	239	10	O82695	O82695 malus domes
14	683	56.1	248	10	Q38742	Q38742 antirrhinum
15	680	55.8	240	10	Q9AR13	Q9ar13 pisum sativ
16	676.5	55.5	245	10	Q9ATE2	Q9ate2 petunia hyb
17	676	55.5	242	10	Q9XHR8	Q9xhr8 nicotiana t
18	670	55.0	242	10	Q9ZTY2	Q9zty2 nicotiana t
19	669	54.9	260	10	Q39400	Q39400 betula verr
20	663	54.4	244	10	Q8S4L5	Q8s415 lycopersico
21	662	54.4	251	10	Q9SEG1	Q9seg1 arabidopsis
22	661	54.3	243	10	Q39399	Q39399 betula verr
23	659	54.1	254	10	Q9XEL0	Q9xel0 sinapis alb
24	656	53.9	242	10	Q9XHR6	Q9xhr6 nicotiana s
25	656	53.9	256	10	Q96355	Q96355 brassica ol
26	654	53.7	256	10	Q96356	Q96356 brassica ol
27	649	53.3	244	10	O82128	O82128 triticum ae
28	648	53.2	248	10	Q8RVR0	Q8rvr0 helianthus
29	648	53.2	256	10	Q39371	Q39371 brassica ol
30	638	52.4	254	10	Q9SBK9	Q9sbk9 brassica ra
31	635	52.1	276	10	Q9LEI0	Q9lei0 hordeum vul
32	634.5	52.1	246	10	Q41355	Q41355 silene lati
33	629	51.6	245	10	Q9ZTT7	Q9ztt7 lolium temu
34	629	51.6	247	10	Q9SNX1	Q9snx1 dendrobium
35	628	51.6	240	10	Q948U1	Q948u1 magnolia pr
36	627	51.5	246	10	Q9MAY7	Q9may7 oryza sativ
37	627	51.5	255	10	Q39081	Q39081 arabidopsis
38	623	51.1	221	10	Q9ZS25	Q9zs25 gerbera hyb
39	621.5	51.0	251	10	Q39375	Q39375 brassica ol
40	621	51.0	246	10	Q9SEX0	Q9sex0 oryza sativ
41	620.5	50.9	261	10	Q9ZTT6	Q9ztt6 lolium temu
42	618	50.7	246	10	Q9M7C6	Q9m7c6 oryza sativ
43	611	50.2	244	10	Q9LEI2	Q9lei2 hordeum vul
44	610.5	50.1	273	10	Q41829	Q41829 zea mays (m
45	607.5	49.9	249	10	Q9SMB3	Q9smb3 oryza sativ

ALIGNMENTS

RESULT 1
Q39401
ID Q39401 PRELIMINARY; PRT; 244 AA.
AC Q39401;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE MADS5 protein.
GN MADS5.
OS Betula verrucosa (White birch) (Betula pendula).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fagales; Betulaceae; Betula.
OX NCBI_TaxID=3505;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=FEMALE INFLORESCENCE;
RA Elo A., Lemmetyinen J., Turunen M.L., Tikka L., Sopanen T.;
RT "Three MADS box genes homologous to SQUAMOSA and APETALA1 have
RT different expression patterns in silver birch (Betula pendula
RT Roth.).";
RL Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR EMBL; X99655; CAA67969.1; -.
DR HSSP; P11746; 1MNM.
DR TRANSFAC; T03048; -.
DR InterPro; IPR002487; TF_Kbox.
DR InterPro; IPR002100; TF_MADSbox.
DR Pfam; PF01486; K-box; 1.
DR Pfam; PF00319; SRF-TF; 1.
DR PRINTS; PR00404; MADSDOMAIN.
DR SMART; SM00432; MADS; 1.
DR PROSITE; PS00350; MADS_BOX_1; 1.
DR PROSITE; PS50066; MADS_BOX_2; 1.
KW DNA-binding; Nuclear protein; Transcription regulation.
SQ SEQUENCE 244 AA; 28174 MW; 358FD25D084ECAB4 CRC64;

Query Match 65.2%; Score 794; DB 10; Length 244;

```
  Best Local Similarity   68.5%;  Pred. No. 1.2e-44;
  Matches  165;  Conservative   28;  Mismatches   42;  Indels    6;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||:||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSTKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| |:|:||:  |: |: :| |||||||||:|||::|:::|:|:|||||||
Db      61 SCMERILERYERYSYADRQLLANDLEQNGSWTLEHAKLKARIEVLQRNQKHFVGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ||||||:|| |||:|:|:| ||||||| |:|||| ||:||||||: || | ||:||:||:
Db     121 SLKELQNLEQQLDSALKHIRSRKNQLMYESISELQRKDKALQEQNNVLAKKVKEKEKELA 180

Qy     181 QQEGQLVQCSNSSSV--LLPQYCVTSSRDGFVERVGGENGGASSLTEP---NSLLPAWML 235
           ||     |     ||  ||||   :|   |  ::  | ||     | |   |:||| |||
Db     181 QQAQWEQQSHTLDSVPSLLPQPLQSSLNIGGSQQARG-NGRVDEGTPPHRANALLPPWML 239

Qy     236 R 236
           |
Db     240 R 240


RESULT 2
Q9XHR7
ID   Q9XHR7      PRELIMINARY;      PRT;   245 AA.
AC   Q9XHR7;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS-box protein MADS1.
GN   NSMADS1.
OS   Nicotiana sylvestris (Wood tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4096;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Jang S., An G.;
RT   "NsMADS1, a member of the MADS gene family from Nicotiana
RT   sylvestris.";
RL   J. Plant Biol. 42:85-87(1999).
DR   EMBL; AF068725; AAD39036.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
SQ   SEQUENCE   245 AA;  28297 MW;  234D53DA0D606450 CRC64;

  Query Match             63.4%;  Score 772.5;  DB 10;  Length 245;
  Best Local Similarity   66.5%;  Pred. No. 3e-43;
  Matches  161;  Conservative   30;  Mismatches   44;  Indels    7;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||| |||||||||||||||| |||||:||||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRASGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| |:::||    |     :| |||||||||:|||::|:|:: |||||||
Db      61 SCMERILERYERYSYAERQLTATDDETPGSWTLEHAKLKARLEVLQRNQRHYAGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |:||||:|||||||:|:| |||||||  | |||| |||||||||: ||:|  |::|||||:
Db     121 SMKELQNLEHQLDSALKHIRSRKNQLMHESISELQKKDKALQEQNNNLSKQVKEREKELA 180

Qy     181 QQ---EGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENG---GASSLTEPNSLLPAWM 234
           ||   | |     |||| :| |  ::|   |      |:||   |:|   : |:::| ||
```

```
Db     181 QQTQWEQQSHDHLNSSSFVLTQ-PLSSLHLGEAYPTAGDNGEVEGSSRQQQQNTVMPPWM 239

Qy     235 LR 236
           ||
Db     240 LR 241


RESULT 3
Q9ZRA5
ID   Q9ZRA5      PRELIMINARY;      PRT;   255 AA.
AC   Q9ZRA5;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS-box protein 2.
OS   Malus domestica (Apple) (Malus sylvestris).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX   NCBI_TaxID=3750;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. FUJI; TISSUE=FLOWER BUDS;
RX   MEDLINE=99373384; PubMed=10444080;
RA   Sung S.K., Yu G.H., An G.;
RT   "Characterization of MdMADS2, a member of the SQUAMOSA subfamily of
RT   genes, in apple.";
RL   Plant Physiol. 120:969-978(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; U78948; AAC83170.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T04707; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   255 AA;  29492 MW;  DA34B2FFD6430203 CRC64;

  Query Match             63.2%;  Score 769.5;  DB 10;  Length 255;
  Best Local Similarity   63.5%;  Pred. No. 5e-43;
  Matches  162;  Conservative   28;  Mismatches   42;  Indels   23;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||||||||:|||||||||||||||:||:||||||||| |
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLMKKAHEISVLCDAEVALIIFSTKGKLFEYSND 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| |:::||:  |   : :| |||||||||||||::|:|:::|||| ||
Db      61 SCMERILERYERYSYTERQLLANDNESTGSWTLEHAKLKARVEVLQRNQRHYMGEDLQSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ||||||:|| |||:|:| |||||||  |:|||| |||||||||: || | ||:||:|
Db     121 SLKELQNLEQQLDSALKHIRSRKNQVMYESISELQKKDKALQEQNNLLAKKVKEKENAVA 180

Qy     181 QQ---EGQLVQCSNSSSVLLPQ-----------YCVTSSRDGFVERVGGENGGASSLT- 224
           ||   |    |  |||| |||:           |    | :|    : |:|      |
Db     181 QQAQLEHVQEQRLNSSSSLLPRALQSLNFGSGSNYQAIRSSEG----IPGDNQQYGDETP 236

Qy     225 ---EPNSLLPAWMLR 236
               || |||||::|
Db     237 TPHRPNMLLPAWIVR 251


RESULT 4
Q9SEG7
ID   Q9SEG7      PRELIMINARY;      PRT;   247 AA.
AC   Q9SEG7;
```

```
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS box protein.
GN   MADS6.
OS   Capsicum annuum (Bell pepper).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
OX   NCBI_TaxID=4072;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21351847; PubMed=11459226;
RA   Sung S.-K., Moon Y.-H., Chung J.-E., Lee S.-Y., Park H.G., An G.;
RT   "Characterization of MADS box genes from hot pepper.";
RL   Mol. Cells 11:352-359(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Jung J.-Y., Moon Y.-H., Chung J.-E., Sung S.-K., An G.;
RL   Submitted (JUN-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; AF130118; AAF22139.2; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   247 AA;  28669 MW;  C44714D6E862F7C1 CRC64;

  Query Match             62.5%;  Score 761.5;  DB 10;  Length 247;
  Best Local Similarity   66.4%;  Pred. No. 1.6e-42;
  Matches  162;  Conservative   30;  Mismatches   43;  Indels    9;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||:||||||||||||||||||||||||||||||||| |||||||||||||||||
Db       1 MGRGRVQLRRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSSKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| |:::|| ||    :| |||||||||:|||::|:|:: |||||||
Db      61 SCMERILERYERYSYAERQLNATDVETPGSWTLEHAKLKARLEVLQRNQRHYAGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |:||||:|| |||:|:| ||||||| | |||| |||||||||: ||:| |::|||||:
Db     121 SMKELQNLEQQLDSALKHIRSRKNQLMHESISELQKKDKALQEQNNNLSKQMKEREKQLA 180

Qy     181 QQ----EGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENG---GASSLTEPN-SLLPA 232
           ||    | |     |||| || :   ::  | |    |:||   |:|   : | :::|
Db     181 QQHTPWEQQNHDHLNSSSFGLP-HPFNNNHLGEVYPTAGDNGEVEGSSRQQQQNAAVMPP 239

Qy     233 WMLR 236
           ||||
Db     240 WMLR 243


RESULT 5
Q9SBQ0
ID   Q9SBQ0      PRELIMINARY;      PRT;   245 AA.
AC   Q9SBQ0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Floral binding protein 26.
OS   Petunia hybrida (Petunia).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
OX   NCBI_TaxID=4102;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. W115;
RX   MEDLINE=20002442; PubMed=10529428;
RA   Immink R.G.H., Hannapel D.J., Ferrario S., Busscher M., Franken J.,
RA   Lookeren Campagne M.M., Angenent G.C.;
RT   "A petunia MADS box gene involved in the transition from vegetative to
RT   reproductive development.";
RL   Development 126:5117-5126(1999).
DR   EMBL; AF176783; AAF19164.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
SQ   SEQUENCE   245 AA;  28161 MW;  1D5913E2D1D8E49E CRC64;

  Query Match             62.1%;  Score 756.5;  DB 10;  Length 245;
  Best Local Similarity   66.1%;  Pred. No. 3.4e-42;
  Matches  160;  Conservative   29;  Mismatches   46;  Indels    7;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||||||||||:|||||||||||||| |||||:||||||||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLRKAHEISVLCDAEVGLIVFSTKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVG-RDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDS 119
           ||||||||||:|| |:::|| |  |     :| |||||||||:|||::|:::: ||||||
Db      61 SCMERILERYERYSYAERQLSGATDNDTPGSWTLEHAKLKARLEVLQRNQKHYAGEDLDS 120

Qy     120 LSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKT 179
           ||:||||:|| |||:|:| ||||||| | |||| |||||||||: || | |::||||:
Db     121 LSMKELQNLEQQLDSALKQIRSRKNQLMHESISELQKKDKALQEQNNKLSKQVKEREKEL 180

Qy     180 GQQEGQLVQCS--NSSSVLLPQYCVTSSRDGFVERVGGENG---GASSLTEPNSLLPAWM 234
           ||      |    |||| :| |  : |   |      |:||   |:|    ||:::| ||
Db     181 AQQSQWEPQSHDLNSSSFVLSQ-PLNSLHLGEAYPSAGDNGEVEGSSRQQPPNTVMPPWM 239

Qy     235 LR 236
           ||
Db     240 LR 241


RESULT 6
Q9SBQ1
ID   Q9SBQ1      PRELIMINARY;      PRT;   246 AA.
AC   Q9SBQ1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS box transcription factor.
GN   PFG.
OS   Petunia hybrida (Petunia).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Petunia.
OX   NCBI_TaxID=4102;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. W115;
RX   MEDLINE=20002442; PubMed=10529428;
RA   Immink R.G.H., Hannapel D.J., Ferrario S., Busscher M., Franken J.,
RA   Lookeren Campagne M.M., Angenent G.C.;
RT   "A petunia MADS box gene involved in the transition from vegetative to
RT   reproductive development.";
RL   Development 126:5117-5126(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; AF176782; AAF19721.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
```

```
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   246 AA;  28539 MW;  C8713413ACD5DA6B CRC64;

  Query Match             61.8%;  Score 753;  DB 10;  Length 246;
  Best Local Similarity   63.1%;  Pred. No. 5.7e-42;
  Matches  157;  Conservative   34;  Mismatches   38;  Indels   20;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||:|||||||||||||||||||||||||||||||||| |||||:|||||||:||
Db       1 MGRGRVQMKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| |:::||| | |   :| |||||||||:||:::|:|::|||||||
Db      61 SCMERILERYERYSYAERQLVSTDHSSPGSWNLEHAKLKARIEVVQRNQRHYMGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |:|:||:|| |||:::| ||||||| | |||| ||||||:||: || | ||:|||||:
Db     121 SMKDLQNLEQQLDSSLKHIRSRKNQLMHESISELQKKDKSLQEQNNLLSKKVKEREKELA 180

Qy     181 Q-----QEGQLVQCSNSSSVLLPQ--------YCVTSSRDGFVERVGGENGGASSLTEPN 227
           |     |:    : ::||| :|||            |: |       || |||   :|
Db     181 QQTQWEQQNNHHEINSSSSFVLPQPLDSPHLGEAYQSTVD------NGEVEGASQ-QQPA 233

Qy     228 SLLPAWMLR 236
           : :| ||||
Db     234 NTMPPWMLR 242


RESULT 7
Q9ZTY3
ID   Q9ZTY3      PRELIMINARY;      PRT;   245 AA.
AC   Q9ZTY3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   NAP1-1.
GN   NAP1-1.
OS   Nicotiana tabacum (Common tobacco).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
OX   NCBI_TaxID=4097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=XANTHI;
RA   Wu Y.H., Li Q., Zhang J.-S., Li Y.;
RL   Submitted (JUN-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; AF009126; AAD01421.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   245 AA;  28223 MW;  B696BCFC1AFB244C CRC64;

  Query Match             61.7%;  Score 751.5;  DB 10;  Length 245;
  Best Local Similarity   64.9%;  Pred. No. 7.1e-42;
  Matches  157;  Conservative   31;  Mismatches   47;  Indels    7;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           ||||||||||||||||||||||||||||||||||||||||| |||||:|||||||:||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVGLIVFSTKGKLFEYSTD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||||||||:|| |:::||   |     :| |||||||| |:|||::|: :: |||||||
Db      61 SCMERILERYERYSYAERQLTATDHETPGSWTLEHAKLKGRLEVLQRNQGHYAGEDLDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           :|||:||||||:|:| ||||||| | |||| |||||:|||: ||:| |::|| ||:
Db     121 CMKELQNLEHQLDSALKHIRSRKNQLMHESISELQKKDEALQEQNNNLSKQVKEGEKELA 180

Qy     181 QQ---EGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENG---GASSLTEPNSLLPAWM 234
           ||   | |     |||| :| |  ::|   |      |:||   |:|   : |:::| ||
Db     181 QQTQWEQQSHDHLNSSSFVLTQ-PLSSLHLGEAYPTAGDNGELEGSSRQQQQNTVMPPWM 239

Qy     235 LR 236
           ||
Db     240 LR 241


RESULT 8
Q41356
ID   Q41356      PRELIMINARY;      PRT;   257 AA.
AC   Q41356;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   SLM5 protein.
GN   SLM5.
OS   Silene latifolia (White campion).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
OX   NCBI_TaxID=37657;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=INFLORESCENCE MERISTEMS AND DEVELOPING FLOWERS;
RX   MEDLINE=95170282; PubMed=7866023;
RA   Hardenack S., Ye D., Saedler H., Grant S.;
RT   "Comparison of MADS-box gene expression in developing male and female
RT   flowers of the dioecious plant white campion.";
RL   Plant Cell 6:1775-1787(1994).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; X80492; CAA56659.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03177; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   257 AA;  29815 MW;  C8D1A1697AD9DEEF CRC64;

  Query Match             60.8%;  Score 741;  DB 10;  Length 257;
  Best Local Similarity   59.8%;  Pred. No. 3.7e-41;
  Matches  150;  Conservative   37;  Mismatches   48;  Indels   16;  Gaps    3;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||:|||||||||||||||:| |||||:|||||||:||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRTGLLKKAHEISVLCDADVGLIVFSTKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:|||||:|| |:::||   |     :| |||||||||:|:|:|| |::|||||:|
Db      61 SCMEKILERYERYSYAERQLTAPDPDSHVSWTLEHAKLKARLEILQKNHRHYMGEDLDTL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKK-- 178
           |||||||: ||||| |:| |||:||| |:|||  |||||||||:|||:| ||:||:||:
```

```
Db    121 SLKELQNFEHQLDTALKHIRSKKNQLMYESIHELQKKDKALQEHNNTLSKKVKEKEKEKE 180

Qy    179 ----TGQQEGQLVQCSNS-------SSVLLPQYCVTSSRD---GFVERVGGENGGASSLT 224
              |: | || :         || :||   :  : :     :     :    |: :
Db    181 KEKAIADQQAQWVQQNQELNPSAFFSSQVLPALNIRGNYETGSSVIVNEVAQTQNRSNSS 240

Qy    225 EPNSLLPAWML 235
            |||:|:|||
Db    241 NNNSLVPSWML 251


RESULT 9
Q9FUI3
ID   Q9FUI3      PRELIMINARY;      PRT;   244 AA.
AC   Q9FUI3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS-box protein EAP1.
OS   Eucalyptus globulus (Blue gum).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Myrtales; Myrtaceae; Eucalyptus.
OX   NCBI_TaxID=34317;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98009992; PubMed=9349279;
RA   Kyozuka J., Harcourt R., Peacock W.J., Dennis E.S.;
RT   "Eucalyptus has functional equivalents of the Arabidopsis AP1 gene.";
RL   Plant Mol. Biol. 35:573-584(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kyozuka J., Harcourt R., Peacock W.J., Dennis E.S.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; AF305076; AAG24909.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   244 AA;  28240 MW;  57AB7102DC01D5B9 CRC64;

  Query Match             60.4%;  Score 735.5;  DB 10;  Length 244;
  Best Local Similarity   64.4%;  Pred. No. 7.9e-41;
  Matches  154;  Conservative   34;  Mismatches   44;  Indels    7;  Gaps    5;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          |||||:|||||||||||||:||||||:|||||||||||||||||||||||:||||||||||
Db      1 MGRGRLQLKRIENKINRQITFSKRRAGLLKKAHEISVLCDAEVALIVFSAKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||||||||:|| |:: |::  :     :| |||||||||:||| :| |:|||||||||
Db     61 SCMERILERYERYSYAEHQVLASETESIGSWTLEHAKLKARLEVLHRNYRHFMGEDLDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          |||:||:|| ||::|:| ||||||| | |||| |||||:|||: || | |||||:|:
Db    121 SLKDLQNLEQQLESALKHIRSRKNQLMHESISVLQKKDRALQEQNNLLTKKIKEKERALA 180

Qy    181 QQEGQLVQCSNS--SSVLLPQYCVTSSRDG-FVERVGGENGGASSLTEP--NSLLPAWM 234
          ||  |  |  ::  | |:|| |  :   :| :  |  | : |  :||:|   :||| |:
Db    181 QQ-AQWEQQDHALDSPVVLPHYLPSLDINGSYQARHNGHDDG-ENLTQPRAGTLLPPWI 237


RESULT 10
Q9FUH8
ID   Q9FUH8      PRELIMINARY;      PRT;   245 AA.
AC   Q9FUH8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS box protein AP2L.
OS   Eucalyptus globulus (Blue gum).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Myrtales; Myrtaceae; Eucalyptus.
OX   NCBI_TaxID=34317;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98009992; PubMed=9349279;
RA   Kyozuka J., Harcourt R., Peacock W.J., Dennis E.S.;
RT   "Eucalyptus has functional equivalents of the Arabidopsis AP1 gene.";
RL   Plant Mol. Biol. 35:573-584(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kyozuka J., Harcourt R., Peacock W.J., Dennis E.S.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; AF306349; AAG30923.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   245 AA;  27991 MW;  C0817603C9A9AB6C CRC64;

  Query Match             60.1%;  Score 732.5;  DB 10;  Length 245;
  Best Local Similarity   63.4%;  Pred. No. 1.2e-40;
  Matches  156;  Conservative   27;  Mismatches   46;  Indels   17;  Gaps    4;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          |||||||||||||||||||||||| ||||||||||||||:|||||||:|||||||:||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRCGLLKKAHEISVLCDADVALIVFSTKGKLFEYATD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          |||||||||:|| |:: |::  :   : || |||||||||:|:|:||::| |||:||||
Db     61 CCMERILERYERYSYAESQVLTNNAETNGNWTLEHAKLKARMEILQKNQKNLMGEELDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||:|||||| |:|:|||||| | | |||| ||:|||||||: || | ||:||:||
Db    121 SLKELQNLEHQLDTALKNIRSRKIQLMCESISELQRKDKALQEQNNMLAKKVKEKEKALA 180

Qy    181 QQEGQLVQCSN-------SSSVLLPQYCVTSSRDGFVERVGGENGGASSLTEPN----SL 229
          ||      |  |       ||||:| |      :  |     | | | |  |  |     :|
Db    181 QQ----TQWDNPQDDGLTSSSVILSQSLQPLNIGGPYHPSGIEEGAA--LGPPQHRNATL 234

Qy    230 LPAWML 235
            |:|||
Db    235 FPSWML 240


RESULT 11
Q9FUI2
ID   Q9FUI2      PRELIMINARY;      PRT;   205 AA.
AC   Q9FUI2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS-box protein EAP2S.
OS   Eucalyptus globulus (Blue gum).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Myrtales; Myrtaceae; Eucalyptus.
OX   NCBI_TaxID=34317;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98009992; PubMed=9349279;
RA   Kyozuka J., Harcourt R., Peacock W.J., Dennis E.S.;
RT   "Eucalyptus has functional equivalents of the Arabidopsis AP1 gene.";
RL   Plant Mol. Biol. 35:573-584(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Kyozuka J., Harcourt R., Peacock W.J., Dennis E.S.;
RL   Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; AF305696; AAG27459.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   205 AA;  23690 MW;  2B1A305864269461 CRC64;

  Query Match             57.8%;  Score 703.5;  DB 10;  Length 205;
  Best Local Similarity   69.8%;  Pred. No. 8e-39;
  Matches  143;  Conservative   23;  Mismatches   28;  Indels   11;  Gaps    2;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||| |||||||||||||||:|||||||:||||||||:||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRCGLLKKAHEISVLCDADVALIVFSTKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           |||||||||:|| |:: |::  :   : || |||||||||:|:|:||::| |||:||||
Db      61 CCMERILERYERYSYAESQVLTNNAETNGNWTLEHAKLKARMEILQKNQKNLMGEELDSL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           ||||||:|||||| |:|:||||| | | |||| ||:|||||: || | ||:||:||
Db     121 SLKELQNLEHQLDTALKNIRSRKIQLMCESISELQRKDKALQEQNNMLAKKVKEKEKALA 180

Qy     181 QQEGQLVQCSN-------SSSVLLP 198
           ||     |  |       |||  ||
Db     181 QQ----TQWDNPQDDGLTSSSGSLP 201


RESULT 12
Q941M9
ID   Q941M9      PRELIMINARY;      PRT;   252 AA.
AC   Q941M9;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS-box transcription factor DEFH28.
OS   Antirrhinum majus (Garden snapdragon).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Lamiales; Veronicaceae; Antirrhinum.
OX   NCBI_TaxID=4151;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21580397; PubMed=11722760;
RA   Mueller B.M., Saedler H., Zachgo S.;
RT   "The MADS-box gene DEFH28 from Antirrhinum is involved in the
RT   regulation of floral meristem identity and fruit development.";
RL   Plant J. 28:169-180(2001).
DR   EMBL; AY040247; AAK72467.1; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PROSITE; PS00350; MADS_BOX_1; UNKNOWN_1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
SQ   SEQUENCE   252 AA;  29293 MW;  20F75CB473836104 CRC64;

  Query Match             57.1%;  Score 695;  DB 10;  Length 252;
  Best Local Similarity   59.0%;  Pred. No. 3.6e-38;
  Matches  147;  Conservative   37;  Mismatches   51;  Indels   14;  Gaps    4;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||:|||||||||||||||||||||||||||||||:|||||||||:
Db       1 MGRGRVQLKRIENKISRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSTKGKLFEYSTE 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           | ||||||||:|| |::|:|        ||| ||: || ||:|:||:| ||::|||| |
Db      61 SSMERILERYERYSYAEKKLTSDSHEPEENWCLEYPKLVARIELLERNIRNYVGEDLDHL 120

Qy     121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
           |::|||||| ||| |:|  |:|||| | |||| ||||:::||| || | ||||: ||:
Db     121 SMRELQSLEQQLDTALKRTRTRKNQLMHESISQLQKKERSLQDQNNILAKKIKDNEKQQN 180

Qy     181 QQEGQL----VQCSNSSSVLLPQ--------YCVTSSRDGFVERVGGENGG-ASSLTEPN 227
           :::  :     | |:| ::||           |:  |  || : |  | ||  : : :
Db     181 EKQQDVHEGFAQSSSSINMLLQPPATQLHAVPCLPIS-GGFQQTVRVEEGGDRTRIADSR 239

Qy     228 SLLPAWMLR 236
           | :| |:|:
Db     240 SHIPPWLLQ 248


RESULT 13
O82695
ID   O82695      PRELIMINARY;      PRT;   239 AA.
AC   O82695;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS-box protein.
GN   MADS5.
OS   Malus domestica (Apple) (Malus sylvestris).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Rosales; Rosaceae; Maloideae; Malus.
OX   NCBI_TaxID=3750;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. GRANNY SMITH;
RA   Yao J.L., Dong Y.H., Kvarnheden A., Morris B.A.M.;
RT   "Seven MADS-box genes in apple are expressed in different parts of the
RT   fruit.";
RL   J. Am. Soc. Hortic. Sci. 124:8-13(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; AJ000759; CAA04321.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   239 AA;  27983 MW;  D2FCC7FA35EBC9F5 CRC64;

  Query Match             56.1%;  Score 683;  DB 10;  Length 239;
  Best Local Similarity   67.8%;  Pred. No. 2.1e-37;
  Matches  135;  Conservative   28;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
           |||||||||||||||||||||||:|||||||||||||:|||||||:||||||:||
Db       1 MGRGRVQLKRIENKINRQVTFSKRRTGLLKKAHEISVLCDAQVALIVFSNKGKLFEYATD 60

Qy      61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
           ||||:||||||:|| |:::|||  |     ||  |:::|||: |||::| |:::|||||||
```

```
Db     61 SCMEQILERYERYSYAERQLVEPDFESQGNWTFEYSRLKAKAEVLQRNHRHYLGEDLDSL 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          :|||:|:|| ||| |:| || |||| | |||| ||:| ||:|: || | |||||:||
Db    121 TLKEIQNLEQQLDTALKQIRLRKNQLMNESISELQRKRKAIQEENNLLAKKIKEKEKAAA 180

Qy    181 QQEGQLVQCSNSSSVLLPQ 199
          | : |  :  |    ||||
Db    181 QPQVQNWEQQNHGLDLLPQ 199


RESULT 14
Q38742
ID   Q38742      PRELIMINARY;      PRT;   248 AA.
AC   Q38742;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   SQUA.
GN   SQUAMOSA.
OS   Antirrhinum majus (Garden snapdragon).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; euasterids I; Lamiales; Veronicaceae; Antirrhinum.
OX   NCBI_TaxID=4151;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=T53;
RX   MEDLINE=92224862; PubMed=1563342;
RA   Huijser P., Klein J., Lonnig W.E., Meijer H., Saedler H., Sommer H.;
RT   "Bracteomania, an inflorescence anomaly, is caused by the loss of
RT   function of the MADS-box gene squamosa in Antirrhinum majus.";
RL   EMBO J. 11:1239-1249(1992).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; X63701; CAA45228.1; -.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T03178; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   248 AA;  28835 MW;  2470CDF92D708319 CRC64;

  Query Match             56.1%;  Score 683;  DB 10;  Length 248;
  Best Local Similarity   70.3%;  Pred. No. 2.2e-37;
  Matches  128;  Conservative   32;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          ||||:||||||||||||||||||| ||||||||:||||||||||||||:||||||||||
Db      1 MGRGKVQLKRIENKINRQVTFSKRRGGLLKKAHELSVLCDAEVALIVFSNKGKLFEYSTD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          |||:|||||:|:|| :::::|||  :     || ||::|||||:|:|::| |::||||||:
Db     61 SCMDRILEKYERYSFAERQLVSNEPQSPANWTLEYSKLKARIELLQRNHRHYMGEDLDSM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||:|||| ||| |:|:||:|||| :::||| || |:||:|: |  | |||||:||:
Db    121 SLKEIQSLEQQLDTALKNIRTRKNQLLYDSISELQHKEKAIQEQNTMLAKKIKEKEKEIA 180

Qy    181 QQ 182
          ||
Db    181 QQ 182


RESULT 15
Q9AR13
ID   Q9AR13      PRELIMINARY;      PRT;   240 AA.
AC   Q9AR13;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   MADS-box transcription factor (MADS-box transcription factor
DE   MADS4).
GN   PEAM4.
OS   Pisum sativum (Garden pea).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; Pisum.
OX   NCBI_TaxID=3888;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. VIKTORIA; TISSUE=FLORAL APICES;
RA   Berbel A., Navarro C., Ferrandiz C., Canas L.A., Madueno F.,
RA   Beltran J.P.;
RT   "Analysis of PEAM4, the pea AP1 functional homologue,supports a model
RT   for AP1 controlling both floral meristemidentity and a function in
RT   different plant species.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. VIKTORIA;
RX   MEDLINE=21159587; PubMed=11260500;
RA   Berbel A., Navarro C., Ferrandiz C., Canas L.A., Madueno F.,
RA   Beltran J.P.;
RT   "PEAM4, the pea functional homologue to APETALA1 and SQUAMOSA, works
RT   as an A-function gene.";
RL   Plant J. 25:441-451(2001).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Taylor S., Hofer J., Murfet I., Sollinger J., Singer S., Knox M.,
RA   Ellis N.;
RT   "Proliferating inflorescence meristem, a MADS-box gene that regulates
RT   floral meristem identity in pea.";
RL   Submitted (DEC-2001) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION FACTORS.
DR   EMBL; AJ279089; CAC37031.1; -.
DR   EMBL; AJ291298; CAC35027.1; -.
DR   EMBL; AF461740; AAL66379.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription regulation.
SQ   SEQUENCE   240 AA;  28184 MW;  653EA580D5196216 CRC64;

  Query Match             55.8%;  Score 680;  DB 10;  Length 240;
  Best Local Similarity   69.7%;  Pred. No. 3.3e-37;
  Matches  129;  Conservative   30;  Mismatches   26;  Indels    0;  Gaps    0;

Qy      1 MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVALIVFSSKGKLFEYSTD 60
          |||||||||||||||||||||||:|||||||||||||||||||||| |||||||:||
Db      1 MGRGRVQLKRIENKINRQVTFSKRRAGLLKKAHEISVLCDAEVALIVFSHKGKLFEYATD 60

Qy     61 SCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARVEVLEKNKRNFMGEDLDSL 120
          ||||:|||||:|| |:::|||  |     || :|: :|||::::|::| |::||||| ::
Db     61 SCMEKILERYERYSYAERQLVANDSESQGNWTIEYTRLKAKIDLLQRNYRHYMGEDLGTM 120

Qy    121 SLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKALQDHNNSLLKKIKEREKKTG 180
          ||||||||| ||| |:| ||:|:|| |:|||| ||||:| :|: || | |||||:||
Db    121 SLKELQSLEQQLDTALKLIRTRRNQLMYESISELQKKEKVIQEQNNMLAKKIKEKEKIAA 180

Qy    181 QQEGQ 185
          :|: |
Db    181 EQQVQ 185
```

Search completed: January 27, 2003, 12:39:02
Job time : 34 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: February 2, 2003, 05:39:47 ; Search time 3036 Seconds
(without alignments)
10180.230 Million cell updates/sec

Title: US-09-981-087A-1
Perfect score: 1062
Sequence: 1 cccagagagacataagaaag.....taatgatttctcatgaaata 1062

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1062	100.0	1062	6	AR138344	AR138344 Sequence
2	1062	100.0	1062	6	AR150973	AR150973 Sequence
3	1058	99.6	1058	8	ATU33473	U33473 Arabidopsis
4	1018	95.9	1042	8	AF386929	AF386929 Arabidops
5	874	82.3	874	8	AY072463	AY072463 Arabidops
6	710.8	66.9	946	8	SAU25695	U25695 Sinapis alb
c 7	345	32.5	61634	8	AB008269	AB008269 Arabidops
8	339.8	32.0	966	8	BPMADS5GN	X99655 B.pendula m
9	322.4	30.4	1230	8	MDU78948	U78948 Malus domes
10	319.6	30.1	970	8	SLSLM5	X80492 S.latifolia
11	312.4	29.4	1085	8	AF305076	AF305076 Eucalyptu
12	311.6	29.3	956	8	AF130118	AF130118 Capsicum
13	311.2	29.3	1032	8	AF305696	AF305696 Eucalyptu
14	311.2	29.3	1070	8	AF306349	AF306349 Eucalyptu
15	310.2	29.2	1029	8	AF068725	AF068725 Nicotiana
16	309.8	29.2	1131	6	A81451	A81451 Sequence 1
17	309.8	29.2	1131	8	AF176782	AF176782 Petunia x
18	309	29.1	1140	8	PSA279089	AJ279089 Pisum sat
19	309	29.1	1207	8	AF461740	AF461740 Pisum sat
20	307.4	28.9	1062	8	STU23758	U23758 Solanum tub
21	307.2	28.9	962	8	LETDR4	X60757 L.esculentu
22	305.8	28.8	1068	8	STU23757	U23757 Solanum tub
23	303.6	28.6	1023	8	SAAP1	X81480 S.alba ap1
24	303.2	28.5	998	8	BPMADS4GN	X99654 B.pendula m
25	302.6	28.5	1079	8	SCAF002666	AF002666 Solanum c
26	301	28.3	794	6	AR042840	AR042840 Sequence
27	301	28.3	794	6	AR095091	AR095091 Sequence
28	301	28.3	794	6	AR111348	AR111348 Sequence
29	301	28.3	794	6	AR200415	AR200415 Sequence
30	301	28.3	1138	8	BOU67452	U67452 Brassica ol
31	300.8	28.3	1143	8	AF009126	AF009126 Nicotiana
32	298.2	28.1	1050	8	AF109403	AF109403 Sinapis a
33	297.8	28.0	1051	8	BOBOAP1MR	Z37968 B.oleracea
34	296.8	27.9	1054	6	AR063254	AR063254 Sequence
35	296.8	27.9	1054	8	ATAP1	Z16421 A.thaliana
36	296.8	27.9	1057	6	AR095090	AR095090 Sequence
37	296.8	27.9	1057	6	AR200414	AR200414 Sequence
38	296.8	27.9	1165	8	AY087956	AY087956 Arabidops
39	296.8	27.9	1215	6	AR042839	AR042839 Sequence
40	296.8	27.9	1215	6	AR111347	AR111347 Sequence
41	291.4	27.4	768	6	AR042841	AR042841 Sequence
42	291.4	27.4	768	6	AR095092	AR095092 Sequence
43	291.4	27.4	768	6	AR111349	AR111349 Sequence
44	291.4	27.4	768	6	AR200416	AR200416 Sequence
45	289.8	27.3	1142	8	BOU67451	U67451 Brassica ol

ALIGNMENTS

RESULT 1
AR138344
LOCUS AR138344 1062 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6198024.
ACCESSION AR138344
VERSION AR138344.1 GI:14479853
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Yanofsky,M.F. and Ferrandiz,C.
TITLE Seed plants characterized by delayed seed dispersal
JOURNAL Patent: US 6198024-A 1 06-MAR-2001;
FEATURES Location/Qualifiers

```
     source          1. .1062
                     /organism="unknown"
BASE COUNT      327 a    193 c    240 g    302 t
ORIGIN

  Query Match           100.0%;  Score 1062;  DB 6;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-228;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
           |||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 2
AR150973
LOCUS       AR150973               1062 bp    DNA     linear   PAT 08-AUG-2001
DEFINITION  Sequence 1 from patent US 6229068.
ACCESSION   AR150973
VERSION     AR150973.1  GI:15115564
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 1062)
  AUTHORS   Yanofsky,M.F., Martienssen,R., Ferrandiz,C. and Gu,Q.
  TITLE     Method of increasing fruit size in a plant
  JOURNAL   Patent: US 6229068-A 1 08-MAY-2001;
FEATURES             Location/Qualifiers
     source          1. .1062
                     /organism="unknown"
BASE COUNT      327 a    193 c    240 g    302 t
ORIGIN

  Query Match           100.0%;  Score 1062;  DB 6;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-228;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
```

```
Qy     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
           |||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 3
ATU33473
LOCUS       ATU33473                1058 bp    mRNA    linear   PLN 17-APR-1996
DEFINITION  Arabidopsis thaliana agamous-like 8 (AGL8) mRNA, complete cds.
ACCESSION   U33473
VERSION     U33473.1  GI:1004364
KEYWORDS    .
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1058)
  AUTHORS   Mandel,M.A. and Yanofsky,M.F.
  TITLE     The Arabidopsis AGL8 MADS box gene is expressed in inflorescence
            meristems and is negatively regulated by APETALA1
  JOURNAL   Plant Cell 7 (11), 1763-1771 (1995)
  MEDLINE   96093419
   PUBMED   8535133
REFERENCE   2  (bases 1 to 1058)
  AUTHORS   Mandel,M.A. and Yanofsky,M.F.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-AUG-1995) M. Alejandra Mandel, Biology, University of
            California at San Diego, La Jolla, CA 92093-0116, USA
FEATURES             Location/Qualifiers
     source          1. .1058
                     /organism="Arabidopsis thaliana"
                     /variety="ecotype Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /map="1.5 cM above LFY"
     gene            1. .1058
                     /gene="AGL8"
     CDS             101. .829
                     /gene="AGL8"
                     /standard_name="agamous-like gene 8"
                     /note="contains MADS box and K domains"
                     /codon_start=1
                     /product="AGL8"
                     /protein_id="AAA97403.1"
                     /db_xref="GI:1004365"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVA
                     LIVFSSKGKLFEYSTDSCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARV
                     EVLEKNKRNFMGEDLDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKA
                     LQDHNNSLLKKIKEREKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENG
                     GASSLTEPNSLLPAWMLRPTTTNE"
BASE COUNT      324 a    193 c    240 g    301 t
ORIGIN

  Query Match           99.6%;  Score 1058;  DB 8;  Length 1058;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-228;
  Matches 1058;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
```

```
Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGA 1058
           |||||||||||||||||||||||||||||||||||||
Db    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGA 1058


RESULT 4
AF386929
LOCUS       AF386929               1042 bp    mRNA    linear   PLN 14-JUN-2001
DEFINITION  Arabidopsis thaliana floral homeotic protein AGL8 (MSL3.3) mRNA,
            complete cds.
ACCESSION   AF386929
VERSION     AF386929.1  GI:14423383
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 1042)
  AUTHORS   Lam,B., Southwick,A., Karlin-Neumann,G., Nguyen,M., Miranda,M.,
            Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
            Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
            Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
            Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
            Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAY-2001) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     e-mail for correspondence: arab@sequence.stanford.edu

            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis  Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T. , Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Lam,B.,
            Karlin-Neumann,G., Nguyen,M., Southwick,A., Miranda,M., Palm,C.J.,
            Bowser,L., Jones,T., Banh,J ., Chen,H., Cheuk,R., Chung,M.K.,
            Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
            Yamada,K., Ecker,J., Theologis,A. and Davis, R.W.

            Lam,B., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally
            to this work.  Shinozaki,K. (RIKEN GSC) and Davis,R.W.
            (SSP/Stanford) contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .1042
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="V"
                     /clone="RAFL05-11-O24; R16143"
                     /note="This clone is in pBluescript
                     ecotype: Columbia"
     gene            1. .1042
                     /gene="MSL3.3"
     CDS             60. .788
                     /gene="MSL3.3"
                     /codon_start=1
                     /product="floral homeotic protein AGL8"
                     /protein_id="AAK62374.1"
                     /db_xref="GI:14423384"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVA
                     LIVFSSKGKLFEYSTDSCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARV
                     EVLEKNKRNFMGEDLDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKA
                     LQDHNNSLLKKIKEREKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENG
                     GASSLTEPNSLLPAWMLRPTTTNE"
BASE COUNT      312 a    190 c    230 g    310 t
ORIGIN

  Query Match             95.9%;  Score 1018;  DB 8;  Length 1042;
  Best Local Similarity   100.0%;  Pred. No. 8.5e-219;
  Matches 1018;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      42 GGTCATTTCAGGGTTGTCGTTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATA 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 GGTCATTTCAGGGTTGTCGTTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATA 60

Qy     102 TGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTT 161
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTT 120

Qy     162 TCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATG 221
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 TCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATG 180

Qy     222 CTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACT 281
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACT 240

Qy     282 CTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTG 341
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTG 300

Qy     342 TTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAA 401
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAA 360

Qy     402 GAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGA 461
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGA 420

Qy     462 GCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGT 521
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGT 480

Qy     522 CAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCT 581
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCT 540

Qy     582 TGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTC 641
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 TGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTC 600

Qy     642 AGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACT 701
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 AGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACT 660

Qy     702 GCGTAACCTCCTCCAGAGATGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCAT 761
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GCGTAACCTCCTCCAGAGATGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCAT 720

Qy     762 CGTCGTTGACGGAACCAAACTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGA 821
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CGTCGTTGACGGAACCAAACTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGA 780

Qy     822 ACGAGTAGAACTATCTCACTCTTTATAATATAATGATAATATAATTAATGTTTAATATTT 881
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 ACGAGTAGAACTATCTCACTCTTTATAATATAATGATAATATAATTAATGTTTAATATTT 840

Qy     882 TCATAACATTCAGCATTTTTTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAG 941
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCATAACATTCAGCATTTTTTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAG 900

Qy     942 CTAGTCATATTATATGTATGATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCA 1001
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 CTAGTCATATTATATGTATGATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCA 960
```

```
Qy   1002 TTAGATTCACTGCGTCTTAAGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAA 1059
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 TTAGATTCACTGCGTCTTAAGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAA 1018


RESULT 5
AY072463
LOCUS       AY072463                 874 bp    mRNA    linear   PLN 27-JAN-2002
DEFINITION  Arabidopsis thaliana floral homeotic protein AGL8 (MSL3.3) mRNA,
            complete cds.
ACCESSION   AY072463
VERSION     AY072463.1  GI:18377423
KEYWORDS    FLI_CDNA.
SOURCE      Arabidopsis thaliana.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 874)
  AUTHORS   Nguyen,M., Karlin-Neumann,G., Southwick,A., Lam,B., Miranda,M.,
            Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H.,
            Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
            Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
            Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K.,
            Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JAN-2002) DNA Sequencing and Technology Center,
            Stanford University, 855 California Avenue, Palo Alto, CA 94304,
            USA
COMMENT     e-mail for correspondence: arab@sequence.stanford.edu

            RIKEN Genomic Sciences Center (GSC) members carried out the
            collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
            Arabidopsis  Full-Length cDNA'): Seki,M., Narusaka,M., Ishida,J.,
            Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
            Hayashizaki,Y. and Shinozaki,K.

            The Salk, Stanford, PGEC (SSP) Consortium members carried out the
            sequencing and annotation of the RAFL cDNAs: Nguyen,M.,
            Southwick,A., Karlin-Neumann,G.,  Lam,B., Miranda,M., Palm,C.J.,
            Bowser,L., Jones,T., Banh,J., Chen,H., Cheuk,R., Chung,M.K.,
            Kim,C., Lin,J., Liu,S.X., Pham,P.K., Sakano,H., Shinn,P.,
            Yamada,K., Ecker,J., Theologis,A. and Davis,R.W.

            Nguyen,M, (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed
            equally to this work.  Shinozaki,K. (RIKEN GSC) and Davis,R.W.
            (SSP/Stanford)  contributed equally to this work as PIs.
FEATURES             Location/Qualifiers
     source          1. .874
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="U16143"
                     /note="This clone is in pENTR/SD-dTopo
                     ecotype: Columbia"
     gene            1. .874
                     /gene="MSL3.3"
     CDS             1. .729
                     /gene="MSL3.3"
                     /codon_start=1
                     /product="floral homeotic protein AGL8"
                     /protein_id="AAL66878.1"
                     /db_xref="GI:18377424"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVA
                     LIVFSSKGKLFEYSTDSCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARV
                     EVLEKNKRNFMGEDLDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKA
                     LQDHNNSLLKKIKEREKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENG
                     GASSLTEPNSLLPAWMLRPTTTNE"
BASE COUNT      267 a    166 c    197 g    244 t
ORIGIN

  Query Match             82.3%;  Score 874;  DB 8;  Length 874;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-186;
  Matches  874;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 60

Qy    161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 120

Qy    221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 180

Qy    281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 240

Qy    341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 300

Qy    401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 360

Qy    461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 420

Qy    521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 480

Qy    581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGT 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGT 540

Qy    641 CAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATAC 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 CAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATAC 600

Qy    701 TGCGTAACCTCCTCCAGAGATGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCA 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TGCGTAACCTCCTCCAGAGATGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCA 660

Qy    761 TCGTCGTTGACGGAACCAAACTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACG 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TCGTCGTTGACGGAACCAAACTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACG 720

Qy    821 AACGAGTAGAACTATCTCACTCTTTATAATATAATGATAATATAATTAATGTTTAATATT 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AACGAGTAGAACTATCTCACTCTTTATAATATAATGATAATATAATTAATGTTTAATATT 780

Qy    881 TTCATAACATTCAGCATTTTTTTGGTGACTTATACTCATTATTAATACCGATATGTTTTA 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TTCATAACATTCAGCATTTTTTTGGTGACTTATACTCATTATTAATACCGATATGTTTTA 840

Qy    941 GCTAGTCATATTATATGTATGATGGAACTCCGTT 974
          |||||||||||||||||||||||||||||||||
Db    841 GCTAGTCATATTATATGTATGATGGAACTCCGTT 874


RESULT 6
SAU25695
LOCUS       SAU25695                 946 bp    mRNA    linear   PLN 29-JAN-1997
DEFINITION  Sinapis alba transcription factor SaMADS B mRNA, complete cds.
ACCESSION   U25695
VERSION     U25695.1  GI:1049023
KEYWORDS    .
SOURCE      Sinapis alba.
  ORGANISM  Sinapis alba
```

```
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Sinapis.
REFERENCE   1  (bases 1 to 946)
  AUTHORS   Menzel,G., Apel,K. and Melzer,S.
  TITLE     Identification of two MADS box genes that are expressed in the
            apical meristem of the long-day plant Sinapis alba in transition to
            flowering
  JOURNAL   Plant J. 9 (3), 399-408 (1996)
  MEDLINE   97077349
   PUBMED   8919916
REFERENCE   2  (bases 1 to 946)
  AUTHORS   Menzel,G., Apel,K. and Melzer,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-APR-1995) Siegbert Melzer, Institute of Plant
            Sciences, Federal Institute of Technology, Universitaetstr. 2,
            Zuerich 8092, Switzerland
FEATURES             Location/Qualifiers
     source          1. .946
                     /organism="Sinapis alba"
                     /db_xref="taxon:3728"
                     /tissue_type="inflorescence meristems"
                     /note="gene expressed in apical meristems in transition to
                     flowering"
     CDS             83. .808
                     /note="MADS box protein"
                     /codon_start=1
                     /product="transcription factor SaMADS B"
                     /protein_id="AAB41525.1"
                     /db_xref="GI:1049024"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVA
                     LVIFSSKGKLFEYSTDSCMEKILERYDRYLYSDKQLVGRDISQSENWVLEHAKLKARV
                     EVLEKNKRNFMGEDLDSLSLKELQSLEHQLHAAIKSIRSRKNQAMFESISALQKKDKV
                     LQDHNNALLKKIKEREKNTVHQEVQLIQCSNNSSILQPQYCLTSSRDGFVGRVEGENE
                     GASSLAEPNSLLPAWMLRPTTNE"
     polyA_site      946
                     /note="16 A nucleotides"
BASE COUNT      303 a    176 c    209 g    258 t
ORIGIN
```

Query Match 66.9%; Score 710.8; DB 8; Length 946;
Best Local Similarity 89.3%; Pred. No. 1.1e-149;
Matches 804; Conservative 0; Mismatches 82; Indels 14; Gaps 3;

```
Qy      15 AGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCGTTTCTCTCTCTTGT 74
           | | | | |||||||||||| | |||||| ||||||| |||| ||||||||||||||||
Db       1 AAAGAAATAGAGAGAGAGAGGTCCTTTGGACATTTCAAGGTTATCGTTTCTCTCTCTTGT 60

Qy      75 TCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAG 134
           ||||||||||||||| |||||||    |||||||| ||||||||||||||||||||||||||
Db      61 TCTTGAGATTTTGAGGAGAGAG----ATGGGAAGGGGTAGGGTTCAGCTGAAGAGGATAG 116

Qy     135 AGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAG 194
           | |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     117 AAAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAG 176

Qy     195 CTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAG 254
           ||||||||||||||||||||||||||||||||||||||| || ||||||||||||||||
Db     177 CTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCGTCATCTTCTCTTCCAAAG 236

Qy     255 GCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATC 314
           |||||||||||||||||||||||||||||||||||| | ||||||||||||||||||||
Db     237 GCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAAAAGATACTTGAACGCTATGATC 296

Qy     315 GCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGG 374
           ||||||| |||||||||||||||||||| ||||||||| ||||||| ||||||||||||
Db     297 GCTATTTGTATTCAGACAAACAACTTGTAGGCCGAGACATTTCACAGAGTGAAAATTGGG 356

Qy     375 TTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATT 434
           ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     357 TTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTCGAGAAGAACAAAAGGAATT 416

Qy     435 TTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGC 494
           |||||||||||||||||||||||||||||||||||||||||||| |||||||||||| ||||
Db     417 TTATGGGGGAAGATCTTGATTCGTTGAGCTTAAAGGAGCTCCAAAGCTTGGAGCACCAGC 476

Qy     495 TCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATAT 554
           || |||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     477 TCCATGCTGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATAT 536

Qy     555 CTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGA 614
           | |||||||||||||||||| | |||||||||||||||||| ||||||||||||||||
Db     537 CAGCGCTCCAGAAGAAGGATAAGGTCTTGCAAGATCACAACAATGCGCTTCTCAAAAAGA 596

Qy     615 TTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACT 674
           |||||||||||||||||| |||| ||| ||||||| |||| || ||||||||||||||
Db     597 TTAAGGAGAGGGAGAAGAACACGGTTCATCAAGAAGTACAACTAATCCAATGCTCCAACA 656

Qy     675 CTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGATGGCTTTGTGGAGA 734
           ||||||| |||| | ||| ||||||||| |||||||||||||||||||||||||||  |
Db     657 ATTCTTCAATTCTCCAGCCCCAATACTGCTTAACCTCCTCCAGAGATGGCTTTGTGGGTA 716

Qy     735 GAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTCTCTGCTTCCGG 794
           |||||| |||||||||||  |||||||||||||| |||||||||||||||| ||||| |
Db     717 GAGTTGAGGGAGAGAACGAAGGTGCATCGTCGTTGGCGGAACCAAACTCTCTCCTTCCAG 776

Qy     795 CTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCACTCTTTATAATATAA 854
           |||||||||||| |||   || |||||||| |  |   |  |  |||| |||||||  |
Db     777 CTTGGATGTTACGCCCT---ACAACGAACGAATGAATAGAATTATCTCTATATAATACGA 833

Qy     855 TGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTTTTTGGTGACTTATA 914
           |        ||||| |||||||||| ||||| |||| |   | ||| |||     ||| |||
Db     834 T-------ACATTAACGTTTAATATTCTCATAGCATTTTGGTGTATTTACTGAACTGATA 886
```

```
RESULT 7
AB008269/c
LOCUS       AB008269               61634 bp    DNA     linear   PLN 27-DEC-2000
DEFINITION  Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MSL3.
ACCESSION   AB008269 BA000015
VERSION     AB008269.1  GI:2618603
KEYWORDS    .
SOURCE      Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
            clone:MSL3.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (sites)
  AUTHORS   Nakamura,Y., Sato,S., Kaneko,T., Kotani,H., Asamizu,E., Miyajima,N.
            and Tabata,S.
  TITLE     Structural analysis of Arabidopsis thaliana chromosome 5. III.
            Sequence features of the regions of 1,191,918 bp covered by
            seventeen physically assigned P1 clones
  JOURNAL   DNA Res. 4 (6), 401-414 (1997)
  MEDLINE   98162728
REFERENCE   2  (bases 1 to 61634)
  AUTHORS   Nakamura,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-OCT-1997) Yasukazu Nakamura, Kazusa DNA Research
            Institute, Department of Plant Gene Research; 1532-3, Yana,
            Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
            Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT     Address for correspondence: kaos@kazusa.or.jp
            For the latest information on annotation of this clone, please see
            http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MSL3
            Genes with similarity to proteins in the databases are described in
            'product' or 'note' qualifiers.  Genes that have no significant
            protein similarity are described as 'unknown protein'.
            The software programs used to predict genes include: Grail
            (Informatics Group, Oak Ridge National Laboratory,
            http://compbio.ornl.gov/Grail-1.3/),
            GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
            NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
            Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
            SplicePredictor (Volker Brendel, Stanford University,
```

```
          http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).
          Genes encoding tRNAs are predicted by tRNAscan-SE
          (Sean Eddy, Washington University School of Medicine, St. Louis,
          http://genome.wustl.edu/eddy/tRNAscan-SE/).
          This sequence may not be the entire insert of this clone. It may be
          shorter because we remove overlaps between neighboring submissions.
          The 5' clone is MAE1 and the 3' clone is MAF19.
FEATURES             Location/Qualifiers
     source          1. .61634
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /chromosome="5"
                     /clone="MSL3"
                     /clone_lib="Mitsui P1"
     exon            complement(1. .106)
                     /note="CDS is reported in Acc# AB015472
                     gene_id:MAE1.12
                     unknown protein"
                     /number=2
                     /evidence=not_experimental
     exon            complement(187. .274)
                     /note="CDS is reported in Acc# AB015472
                     gene_id:MAE1.12
                     unknown protein"
                     /number=1
                     /evidence=not_experimental
     CDS             join(5600. .5732,5825. .5954,6167. .6791)
                     /note="gb|AAC16897.1
                     gene_id:MSL3.1"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="ATR1"
                     /protein_id="BAB10639.1"
                     /db_xref="GI:10177313"
                     /translation="MVRTPCCKEEGIKKGAWTPEEDQKLIAYLHLHGEGGWRTLPEKA
                     GLKRCGKSCRLRWANYLRPDIKRGEFSPEEDDTIIKLHALKGNKWAAIATSLAGRTDN
                     EIKNYWNTNLKKRLKQKGIDAITHKPINSTGQTGFEPKVNKPVYSSGSARLLNRVASK
                     YAVELNRDLLTGIISGNSTVAEDSQNSGDVDSPTSTLLNKMAATSVLINTTTTYSGFS
                     DNCSFTDEFNEFFNNEEISDIYTTVDNFGFMEELKSILSYGDASAGVIENSPEVNVAD
                     AMEFIDSWNEDDNMVGVFV"
     CDS             complement(join(9041. .11023,11553. .12065))
                     /note="contains similarity to receptor-like protein
                     kinase
                     gene_id:MSL3.2"
                     /pseudo
                     /codon_start=1
                     /evidence=not_experimental
     CDS             complement(join(13307. .13418,13555. .13658,14367. .14408,
                     14577. .14618,14758. .14857,15278. .15342,15428. .15506,
                     16400. .16584))
                     /note="gene_id:MSL3.3
                     sp|Q38876"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="floral homeotic protein AGL8"
                     /protein_id="BAB10640.1"
                     /db_xref="GI:10177314"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVA
                     LIVFSSKGKLFEYSTDSCMERILERYDRYLYSDKQLVGRDVSQSENWVLEHAKLKARV
                     EVLEKNKRNFMGEDLDSLSLKELQSLEHQLDAAIKSIRSRKNQAMFESISALQKKDKA
                     LQDHNNSLLKKIKEREKKTGQQEGQLVQCSNSSSVLLPQYCVTSSRDGFVERVGGENG
                     GASSLTEPNSLLPAWMLRPTTTNE"
     CDS             complement(join(22037. .22323,22411. .22694,22785. .22944,
                     23050. .23379,23464. .23491))
                     /note="emb|CAA07251.1
                     gene_id:MSL3.4"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="phytochelatin synthetase"
                     /protein_id="BAB10641.1"
                     /db_xref="GI:10177315"
                     /translation="MVGAQTTEQGDCSKYKGNIPHCCKKDPTVVDLLPGTPYNQQIAN
                     CCKGGVMNSWVQDPATAASSFQISVGAAGTTNKTVRVPRNFTLMGPGPGYTCGPAKIV
                     RPTKFVTTDTRRTTQAMMTWNITCTYSQFLAQRTPTCCVSLSSFYNETIVGCPTCACG
                     CQNNRTESGACLDPDTPHLASVVSPPTKKGTVLPPLVQCTRHMCPIRVHWHVKQNYKE
                     YWRVKITITNFNYRLNYTQWNLVAQHPNLDNITQIFSFNYKSLTPYAGLNDTAMLWGV
                     KFYNDFLSEAGPLGNVQSEILFRKDQSTFTFEKGWAFPRRIYFNGDNCVMPPPDSYPF
                     LPNGGSRSQFSFVAAVLLPLLVFFFFSA"
     CDS             complement(join(26143. .27309,27399. .27543,27666. .27893,
                     27989. .28060,28173. .28258,28393. .28418,28502. .28595,
                     28677. .28748,29027. .29107,29196. .29312,29393. .29506,
                     29582. .29680,29771. .29871,30079. .30277,30355. .30438,
                     30521. .30634,30717. .30830,30925. .31016,31117. .31153,
                     31273. .31397,31513. .31688,31904. .32157,32235. .32387,
                     32489. .32635,32972. .33082))
                     /note="gene_id:MSL3.5"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="kinesin-like protein"
                     /protein_id="BAB10642.1"
                     /db_xref="GI:10177316"
                     /translation="MESTECVRVAVNIRPLITPELLNGCTDCITVAPKEPQVHIGSHT
                     FTYDFVYGNGGYPCSEIYNHCVAPLVDALFKGYNATVLAYGQTGSGKTYTMGTNYSGD
                     CTNGGVIPNVMEDIFRRVETTKDSSELLIRVSFIEIFKEEVFDLLDSNSSALLKNDSG
                     VQAKHTALSRAPIQIRETASGGITLAGVTEAEVKTKEEMGSFLARGSLSRATGSTNMN
                     SQSSRSHAIFTITLEQKKIAGGSCTTTEDGGEDILCAKLHLVDLAGSERAKRTGADGM
                     RLKEGIHINKGLLALGNVISALGDEKKRKEGGHVPYRDSKLTRLLQDSLGGNSKTVMI
                     ACVSPADTNAEETLNTLKYANRARNIQNKAVINRDPATAQMQRMRSQIEQLQTELLFY
                     RGDSGAFDELQILKHKISLLEASNRELHNELQERRVASEHFSKRAYDAQVEKDKLIMI
                     IESVRNGKSLDEIESCQNEDVGLVNKYVSKIQELEGELLHIKNLKKTSNHQYSDDSYD
                     VGPRSNNVLFPSSNESSDCEDKVMDVTDELEFQEKEIEHCSLQEKLDMELKELDKRLE
                     EKEAEMKRFSSGGTSVLKQHYEKKVYDLEQEKRALQREIEGLRHNLASIPSGPGDGAQ
                     KLKEEYVQKLNTLETQVSVLKKKQDAQAQLMRQKQKSDDAAIKLQDEIHRIKSQKVQL
                     QQKIKQESEQFRAWKASREKEVMQLKKEGRRNEYEMHKLMALNQKQKLVLQRKTEEAS
                     QVTKRLKELLDNRKASSRETLSGANGPGTQALMQAIEHEIEVTVRVHEVRSEYERQTE
                     ERARMAKEVARLREENELLKNAKISVHGDTMSPGARNSRIFALENMLATSSSTLVSMA
                     SQLSEAEERERVFGGRGRWNQVRTLGDAKSIMNYLFNLASTARCLARDKEADCREKDV
                     LIRDLKEKIVKFSSYVRYMEIQKADLVHQVKAQTSAMKKLSADENLKNEHSMKKQETR
                     NSTIVLEDMDTSDSEASDHEREDPDLDDEWKPEHESERESEQESVIKLNRKRNFKVGR
                     RRSSVVMRRSYEENSETPSDDAVKSDVCCCTCSKSSSCKTMKCQCRATKGSCGPSCGC
                     SSVKCSNRNADGKENNSISESEALENGENSQESDEKDKGQQQQVLASRGAMLLQNALA
                     DKPEEETNDDGGTRRRRKPLSDIGNTTVKNSLKFPLFWSFQHICVHVARKQQLSDLMK
                     FWFQGKSNVPRPSQRKKWKKTVLQLVPVGPPALPPTHTNTHLIPEANSVTVDSDTARM
                     PENSDSGESNSIKLKLPRAMRSASSNGSNLLRERNADQNGSESGGNSGFVQSNSGRAS
                     GSRTSDEKENHTRRV"
     CDS             join(33663. .33767,33983. .34182,34351. .34420,34537. .34620,
                     34737. .34796,34884. .34954,35056. .35107,35197. .35254,
                     35371. .35471,35571. .35725,35821. .35929,36014. .36226)
                     /note="gene_id:MSL3.6"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="cleavage stimulation factor subunit 1-like
                     protein"
                     /protein_id="BAB10643.1"
                     /db_xref="GI:10177317"
                     /translation="MGNSGDLEQALQDGNIFRQLNALIVAHLRHHNLSQVASAVASAT
                     MTPLNIEVPPNRLLELVAKVIVLHPLKFLYIFLESVKKCFGSEMRTGSCSGEQWDIES
                     VNHAKGSSKTIPKHESKTLSEHKSVVRCARFSPDGMFFATGGADTSIKLFEMISGDTQ
                     ARPLIRTFYDHAEPINDLDFHPRSTILISSAKDNCIKFFDFSKTTAKRAFKVFQDTHN
                     VRSISFHPSGEFLLAGTDHPIPHLYDVNTYQCFLPSNFPDSGVSGAINQVRYSSTGSI
                     YITASKDGAIRLFDGVSAKCVRSIGNAHGKSEVTSAVFTKDQRFVLSSGKDSTVKLWE
                     IGSGRMVKEYLGAKRVKLRSQAIFNDTEEFVISIDEASNEVVTWDARTADKVAKWPSN
                     HNGAPRWIEHSPVESVFVTCGIDRSIRFWKESV"
     CDS             complement(join(37728. .37768,37854. .38286,38358. .38431,
                     38510. .38576))
                     /note="contains similarity to phytochelatin synthetase
                     gene_id:MSL3.7"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="BAB10644.1"
                     /db_xref="GI:10177318"
                     /translation="MESLFSTMIVLLLVSFSCLISTEALTSNYGNITVKWDLLNWTPD
                     GYVAVVTAYNYQKRSIPGWKMSWRGTKKEVIWNMLGAKTTGQGGCSMFKGNIPQSCV
                     RKPTVVDLLPGTPFNQQIANCCKSGVLKPGSESAFQLSVGSAGNSVKTARMPANFMFT
                     APKQQYICGPSKNVRPTRFTTADKRRITAALMTWNITCVFHKAT"
```

```
     CDS             complement(38994. .40559)
                     /note="gb|AAF26800.1
                     gene_id:MSL3.8
                     similar to unknown protein"
                     /codon_start=1
                     /evidence=not_experimental
                     /protein_id="BAB10645.1"
                     /db_xref="GI:10177319"
                     /translation="MPPSLPSLQLRRLLLRSFISSSSVNTLQSQPRIISSKPLFSPLP
                     PSRSSIFSTFPSRFFSSETNAESESLDSNEIALSFSKELTGNPDAESQTISQRFNLSF
                     SHITPNPDLILQTLNLSPEAGRAALGFNEWLDSNSNFSHTDETVSFFVDYFGRRKDFK
                     GMLEIISKYKGIAGGKTLESAIDRLVRAGRPKQVTDFFEKMENDYGLKRDKESLTLVV
                     KKLCEKGHASIAEKMVKNTANEIFPDENICDLLISGWCIAEKLDEATRLAGEMSRGGF
                     EIGTKAYNMMLDCVCKLCRKKDPFKLQPEVEKVLLEMEFRGVPRNTETFNVLINNLCK
                     IRRTEEAMTLFGRMGEWGCQPDAETYLVLIRSLYQAARIGEGDEMIDKMKSAGYGELL

  Query Match             32.5%;  Score 345;  DB 8;  Length 61634;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-67;
  Matches  345;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     715 CAGAGATGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGA 774
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13421 CAGAGATGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGA 13362

Qy     775 ACCAAACTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTA 834
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13361 ACCAAACTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTA 13302

Qy     835 TCTCACTCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAG 894
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13301 TCTCACTCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAG 13242

Qy     895 CATTTTTTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTAT 954
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13241 CATTTTTTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTAT 13182

Qy     955 ATGTATGATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGC 1014
         |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   13181 ATGTATGATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGC 13122

Qy    1015 GTCTTAAGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAA 1059
         |||||||||||||||||||||||||||||||||||||||||||||
Db   13121 GTCTTAAGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAA 13077


RESULT 8
BPMADS5GN
LOCUS       BPMADS5GN               966 bp    mRNA    linear   PLN 20-AUG-1996
DEFINITION  B.pendula mRNA for MADS box protein, MADS5.
ACCESSION   X99655
VERSION     X99655.1  GI:1483231
KEYWORDS    MADS box protein; MADS5 gene.
SOURCE      European white birch.
  ORGANISM  Betula pendula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fagales; Betulaceae; Betula.
REFERENCE   1  (bases 1 to 966)
  AUTHORS   Elo,A., Lemmetyinen,J., Turunen,M.L., Tikka,L. and Sopanen,T.
  TITLE     Three MADS box genes homologous to SQUAMOSA and APETALA1 have
            different expression patterns in silver birch (Betula pendula
            Roth.)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 966)
  AUTHORS   Elo,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-JUL-1996) A. Elo, University Of Joensuu, Department
            Of Biology, P.O. Box 111, 80101 Joensuu, FINLAND
FEATURES             Location/Qualifiers
     source          1. .966
                     /organism="Betula pendula"
                     /db_xref="taxon:3505"
                     /tissue_type="female inflorescence"
     gene            74. .808
                     /gene="MADS5"
     CDS             74. .808
                     /gene="MADS5"
                     /codon_start=1
                     /product="MADS5 protein"
                     /protein_id="CAA67969.1"
                     /db_xref="GI:1483232"
                     /db_xref="SPTREMBL:Q39401"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVA
                     LIVFSTKGKLFEYSTDSCMERILERYERYSYADRQLLANDLEQNGSWTLEHAKLKARI
                     EVLQRNQKHFVGEDLDSLSLKELQNLEQQLDSALKHIRSRKNQLMYESISELQRKDKA
                     LQEQNNVLAKKVKEKEKELAQQAQWEQQSHTLDSVPSLLPQPLQSSLNIGGSQQARGN
                     GRVDEGTPPHRANALLPPWMLRHLNQ"
BASE COUNT      297 a    181 c    217 g    271 t
ORIGIN

  Query Match             32.0%;  Score 339.8;  DB 8;  Length 966;
  Best Local Similarity   75.5%;  Pred. No. 3.5e-66;
  Matches  422;  Conservative    0;  Mismatches  137;  Indels    0;  Gaps    0;

Qy      89 AGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAAT 148
           | | | | |  |||||| ||||| ||||||||| |||| ||||| ||||||||||||||
Db      62 ATATATATAATTATGGGGAGAGGAAGGGTTCAGTTGAAAAGGATCGAGAACAAGATCAAC 121

Qy     149 AGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCT 208
           ||||| || |||||||| |||||| |||| |||||||| |||||||| ||||||||||||
Db     122 AGGCAGGTCACTTTCTCCAAGAGACGGTCAGGTTTGCTGAAGAAAGCCCATGAGATCTCT 181

Qy     209 GTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAA 268
           || || || |||||||||| ||  | || |||||||  ||||||| || ||||| ||
Db     182 GTGCTTTGTGATGCTGAGGTGGCCTTGATTGTCTTCTCCACCAAAGGGAAGCTCTTTGAG 241

Qy     269 TATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCA 328
           ||||| || || ||||| ||||| ||||| |||||||| |||||  | |||| |||| ||
Db     242 TATTCTACTGATTCTTGTATGGAAAGGATCCTTGAACGGTATGAGAGATATTCATATGCA 301

Qy     329 GACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCT 388
           ||||  || ||| |||     ||  ||  ||||| || ||  |||  ||| |||||||||
Db     302 GACAGGCAGCTTCTTGCAAATGATCTTGAACAAAATGGAAGCTGGACTCTGGAACATGCA 361

Qy     389 AAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGAT 448
           |||||||||| ||  |||||||  |  | |  ||  | | | ||||  |||| ||||||
Db     362 AAGCTCAAGGCTAGGATTGAGGTTTTGCAAAGAAATCAGAAGCATTTCGTGGGAGAAGAT 421

Qy     449 CTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATC 508
           || ||||| |||||  | || ||||| ||||  |||||||| ||||| ||| | ||| ||
Db     422 CTCGATTCCTTGAGTCTTAAAGAGCTTCAAAATTTGGAGCAACAGCTTGATTCTGCTCTC 481

Qy     509 AAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAG 568
           ||   ||| ||||||||||||||||||  ||||| |||||| || || | ||| ||||
Db     482 AAACACATCAGGTCAAGAAAGAACCAACTTATGTACGAATCGATTTCAGAGCTTCAGAGA 541

Qy     569 AAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAG 628
           |||||||| || ||||| || || |||||   ||||   || ||| | ||||| | |||
Db     542 AAGGATAAGGCATTGCAGGAGCAAAACAACGTGCTTGCAAAGAAGGTGAAGGAAAAGGAA 601

Qy     629 AAGAAAACGGGTCAGCAAG 647
           ||| ||   |  || || |
Db     602 AAGGAATTAGCCCAACAGG 620


RESULT 9
MDU78948
LOCUS       MDU78948               1230 bp    mRNA    linear   PLN 26-AUG-1999
DEFINITION  Malus domestica MADS-box protein 2 mRNA, complete cds.
ACCESSION   U78948
VERSION     U78948.1  GI:3947984
KEYWORDS    .
SOURCE      Malus x domestica.
  ORGANISM  Malus x domestica
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

```
            Rosidae; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
REFERENCE   1  (bases 1 to 1230)
  AUTHORS   Sung,S.K., Yu,G.H. and An,G.
  TITLE     Characterization of MdMADS2, a member of the SQUAMOSA subfamily of
            genes, in apple
  JOURNAL   Plant Physiol. 120 (4), 969-978 (1999)
  MEDLINE   99373384
   PUBMED   10444080
REFERENCE   2  (bases 1 to 1230)
  AUTHORS   Sung,S.-K. and An,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-NOV-1996) Life Science, Pohang University of Science
            and Technology, San 31 Hyoja Dong, Nam-gu, Pohang 790-784, Republic
            of Korea
FEATURES             Location/Qualifiers
     source          1. .1230
                     /organism="Malus x domestica"
                     /cultivar="Fuji"
                     /db_xref="taxon:3750"
                     /tissue_type="floral bud"
     CDS             234. .1001
                     /function="floral development"
                     /function="flowering time"
                     /note="MdMADS2; transcription factor"
                     /codon_start=1
                     /product="MADS-box protein 2"
                     /protein_id="AAC83170.1"
                     /db_xref="GI:3947985"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRSGLMKKAHEISVLCDAEVA
                     LIIFSTKGKLFEYSNDSCMERILERYERYSYTERQLLANDNESTGSWTLEHAKLKARV
                     EVLQRNQRHYMGEDLQSLSLKELQNLEQQLDSALKHIRSRKNQVMYESISELQKKDKA
                     LQEQNNLLAKKVKEKENAVAQQAQLEHVQEQRLNSSSSLLPRALQSLNFGSGSNYQAI
                     RSSEGIPGDNQQYGDETPTPHRPNMLLPAWIVRHLNE"
BASE COUNT      406 a    223 c    274 g    327 t
ORIGIN

  Query Match           30.4%;  Score 322.4;  DB 8;  Length 1230;
  Best Local Similarity  74.3%;  Pred. No. 2.8e-62;
  Matches  407;  Conservative    0;  Mismatches  141;  Indels    0;  Gaps    0;

Qy     100 TATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTAC 159
           |||||| || || ||||| |||||||||||| || |||||||||||||| ||||| || ||
Db     233 TATGGGGAGGGGAAGGGTGCAGCTGAAGAGAATTGAGAACAAGATCAACAGGCAGGTGAC 292

Qy     160 TTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGA 219
           |||||||||||||||||| ||  || | |||||||||||||||||| ||||| || || ||
Db     293 CTTCTCAAAGAGAAGGTCGGGGCTGATGAAGAAAGCTCATGAGATTTCTGTGCTTTGTGA 352

Qy     220 TGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGA 279
           |||||||||||| | ||  |||||||  |||| ||||| ||||| || || |||| ||
Db     353 TGCTGAGGTTGCTTTGATTATCTTCTCCACCAAGGGCAAGCTCTTTGAGTACTCCAATGA 412

Qy     280 CTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACT 339
           || |||||||| ||||| || ||| | || ||  | || | |||| |||| |  || ||
Db     413 TTCCTGCATGGAAAGGATCCTGGAAAGGTACGAAAGATACTCATATACAGAGAGGCAGCT 472

Qy     340 TGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGC 399
           | |||     ||   |  |   | || ||  |||  ||| |||||||| ||||||||||
Db     473 TCTTGCAAATGATAATGAATCCACTGGAAGCTGGACTCTGGAACATGCAAAGCTCAAGGC 532

Qy     400 AAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTT 459
           || || |||||  |  | |  ||  ||||  | | |||||| ||||||||  | || ||
Db     533 TAGGGTGGAGGTTTTACAAAGAAATCAAAGACACTATATGGGAGAAGATCTCCAATCCTT 592

Qy     460 GAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAG 519
           ||  | || ||||| |||| || ||||| ||||| ||| | ||  | |||  ||| ||
Db     593 AAGTCTCAAAGAGCTTCAAAATTTAGAGCAACAGCTTGATTCTGCACTGAAGCACATAAG 652

Qy     520 GTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGC 579
           |||||||||||||||| ||||| |||||| || |||| ||||||||||||||||||| ||
Db     653 GTCAAGAAAGAACCAAGTTATGTACGAATCGATTTCTGAGCTCCAGAAGAAGGATAAGGC 712

Qy     580 CTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGG 639
           |||||| || || ||||| | |||    || ||| | ||||||| ||||||   |   |
Db     713 ATTGCAGGAGCAAAACAACTTGCTGGCAAAGAAGGTGAAGGAGAAGGAGAACGCAGTAGC 772

Qy     640 TCAGCAAG 647
           ||| || |
Db     773 TCAACAGG 780


RESULT 10
SLSLM5
LOCUS       SLSLM5                   970 bp    mRNA    linear   PLN 18-DEC-1996
DEFINITION  S.latifolia SLM5 mRNA.
ACCESSION   X80492
VERSION     X80492.1  GI:602907
KEYWORDS    MADS-box protein; SLM5 gene.
SOURCE      Silene latifolia.
  ORGANISM  Silene latifolia
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Caryophyllidae; Caryophyllales; Caryophyllaceae; Silene.
REFERENCE   1  (bases 1 to 970)
  AUTHORS   Hardenack,S., Ye,D., Saedler,H. and Grant,S.
  TITLE     Comparison of MADS box gene expression in developing male and
            female flowers of the dioecious plant white campion
  JOURNAL   Plant Cell 6 (12), 1775-1787 (1994)
  MEDLINE   95170282
   PUBMED   7866023
REFERENCE   2  (bases 1 to 970)
  AUTHORS   Hardenack,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUL-1994) S. Hardenack, Max Planck Inst fuer
            Zuechtungsforsch., Carl von Linne Weg 10, 5000 Koeln 30, FRG
FEATURES             Location/Qualifiers
     source          1. .970
                     /organism="Silene latifolia"
                     /sub_species="alba"
                     /db_xref="taxon:37657"
                     /tissue_type="inflorescence meristems and developing
                     flowers"
     gene            22. .795
                     /gene="SLM5"
     CDS             22. .795
                     /gene="SLM5"
                     /codon_start=1
                     /protein_id="CAA56659.1"
                     /db_xref="GI:602908"
                     /db_xref="SPTREMBL:Q41356"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRTGLLKKAHEISVLCDADVG
                     LIVFSTKGKLFEYATDSCMEKILERYERYSYAERQLTAPDPDSHVSWTLEHAKLKARL
                     EILQKNHRHYMGEDLDTLSLKELQNFEHQLDTALKHIRSKKNQLMYESIHELQKKDKA
                     LQEHNNTLSKKVKEKEKEKEKEKAIADQQAQWVQQNQELNPSAFFSSQVLPALNIRGN
                     YETGSSVIVNEVAQTQNRSNSSNNNSLVPSWMLNHLAEQ"
BASE COUNT      341 a    177 c    219 g    233 t
ORIGIN

  Query Match           30.1%;  Score 319.6;  DB 8;  Length 970;
  Best Local Similarity  73.8%;  Pred. No. 1.2e-61;
  Matches  406;  Conservative    0;  Mismatches  144;  Indels    0;  Gaps    0;

Qy     101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
           |||||  |||||||||||||| |||||||||| |||||||||||||| | ||||| ||
Db      22 ATGGGGCGAGGTAGGGTTCAACTGAAGAGGATCGAGAACAAGATCAATCGTCAAGTGACG 81

Qy     161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
           ||||| || ||  |  | |||||||| ||||||||||||||||| ||||| || || |||
Db      82 TTCTCGAAAAGGCGAACCGGTTTGCTTAAGAAAGCTCATGAGATATCTGTGCTATGTGAT 141

Qy     221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
           ||||| |||| ||| || || |||||| |||| |||||||||| || ||| | |||||
Db     142 GCTGATGTTGGTCTTATTGTTTTCTCTACCAAGGGCAAACTCTTTGAGTATGCTACCGAT 201

Qy     281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
           || |||||||||| |||||| ||||| ||||| || || | |||| |||| | ||||||
```

```
Db    202 TCATGCATGGAGAAGATACTAGAACGATATGAACGTTACTCATATGCAGAGAGACAACTG 261

Qy    341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
          ||  |  ||   |        ||  |  |||  | | || ||||||||||||||||||
Db    262 ACTGCACCTGATCCTGATTCTCATGTTAGCTGGACTTTGGAGCATGCTAAGCTCAAGGCT 321

Qy    401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
          ||| ||||  | ||  |||||||  | ||| ||| ||||||||||||||| ||| | |||
Db    322 AGACTTGAAATTCTACAGAAGAATCATAGGCATTATATGGGGGAAGATCTAGATACCTTG 381

Qy    461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
          || |||||||||||| || |  || || ||||| || ||  | ||| | ||   |||  |
Db    382 AGTTTGAAGGAGCTTCAGAATTTCGAACATCAACTTGACACTGCTCTTAAACACATCCGA 441

Qy    521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
          || |  ||||||||    |||| |||||| ||    | ||| |||||||||| |||||
Db    442 TCCAAGAAGAACCAGCTCATGTACGAATCAATTCACGAGCTTCAGAAGAAGGACAAAGCA 501

Qy    581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGT 640
          |||||  || || |||||| | ||    || ||| | ||||||| ||||||| | | ||
Db    502 TTGCAGGAGCATAACAATACCCTATCAAAGAAGGTGAAGGAGAAGGAGAAGGAGAAGGAA 561

Qy    641 CAGCAAGAAG 650
          || || | |
Db    562 AAGGAAAAGG 571


RESULT 11
AF305076
LOCUS       AF305076                1085 bp    mRNA    linear   PLN 23-OCT-2000
DEFINITION  Eucalyptus globulus MADS-box protein EAP1 mRNA, complete cds.
ACCESSION   AF305076
VERSION     AF305076.1  GI:10946428
KEYWORDS    .
SOURCE      Eucalyptus globulus.
  ORGANISM  Eucalyptus globulus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Myrtales; Myrtaceae; Eucalyptus.
REFERENCE   1  (bases 1 to 1085)
  AUTHORS   Kyozuka,J., Harcourt,R., Peacock,W.J. and Dennis,E.S.
  TITLE     Eucalyptus has functional equivalents of the Arabidopsis AP1 gene
  JOURNAL   Plant Mol. Biol. 35 (5), 573-584 (1997)
  MEDLINE   98009992
   PUBMED   9349279
REFERENCE   2  (bases 1 to 1085)
  AUTHORS   Kyozuka,J., Harcourt,R., Peacock,W.J. and Dennis,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2000) Plant Industry, CSIRO, GPO Box 1600,
            Canberra, ACT 2601, Australia
FEATURES             Location/Qualifiers
     source          1. .1085
                     /organism="Eucalyptus globulus"
                     /db_xref="taxon:34317"
     CDS             107. .841
                     /codon_start=1
                     /product="MADS-box protein EAP1"
                     /protein_id="AAG24909.1"
                     /db_xref="GI:10946429"
                     /translation="MGRGRLQLKRIENKINRQITFSKRRAGLLKKAHEISVLCDAEVA
                     LIVFSAKGKLFEYSTDSCMERILERYERYSYAEHQVLASETESIGSWTLEHAKLKARL
                     EVLHRNYRHFMGEDLDSLSLKDLQNLEQQLESALKHIRSRKNQLMHESISVLQKKDRA
                     LQEQNNLLTKKIKEKERALAQQAQWEQQDHALDSPVVLPHYLPSLDINGSYQARHNGH
                     DDGENLTQPRAGTLLPPWISASQLRL"
BASE COUNT      298 a    258 c    250 g    279 t
ORIGIN

  Query Match             29.4%;  Score 312.4;  DB 8;  Length 1085;
  Best Local Similarity   70.8%;  Pred. No. 5e-60;
  Matches  415;  Conservative    0;  Mismatches  171;  Indels    0;  Gaps    0;

Qy     62 TTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAG 121
          || |  |   |||  | |    | ||||  |       ||||| || || ||| | |||
```

```
Db     68 TTTTGGCGTGTGTGTTGGGTTGTGTGAATTGGAGCGAGGATGGGGAGGGGGAGGCTGCAG 127

Qy    122 CTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGT 181
          |||||||||||||||||||||||||  | ||| | || ||||| |||| ||| | |||
Db    128 CTGAAGAGGATAGAGAACAAGATCAACCGACAAATCACCTTCTCCAAGAGGAGGGCGGGT 187

Qy    182 TTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTC 241
          ||||||||||  || || |||||||| || |||||||| || ||||| || |||||||||
Db    188 CTGCTCAAGAAGGCCCACGAGATCTCCGTCCTCTGCGACGCCGAGGTCGCCCTCATCGTC 247

Qy    242 TTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACTT 301
          |||||  |||| ||||| |||||||| || |||||||| || ||||||||||| || ||
Db    248 TTCTCCGCCAAGGGCAAGCTCTTCGAGTACTCCACCGATTCCTGCATGGAGAGAATTCTC 307

Qy    302 GAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAA 361
          ||||||||||  | || | |||| | ||  | ||| || |||   | ||      |
Db    308 GAACGCTATGAAAGATACTCATATGCGGAGCACCAAGTTCTTGCAAGCGAGACGGAATCG 367

Qy    362 AGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAG 421
          | ||  |  |||  | | || |||||||||||||||| ||| |||| ||  |  | |
Db    368 ATTGGTAGCTGGACTTTGGAGCATGCTAAGCTCAAGGCCAGACTTGAAGTTTTACACAGA 427

Qy    422 AACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGC 481
          ||  | ||| |||| ||||| |||||||||||||| |||||  | |||||| |||||||
Db    428 AATTATAGGCATTTCATGGGAGAAGATCTTGATTCTTTGAGTCTCAAGGACCTCCAAAAT 487

Qy    482 TTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATG 541
          |||||||| || || ||  | ||| | ||   ||| || || |||||||| ||    |||
Db    488 TTGGAGCAGCAACTGGAGTCTGCTCTTAAACACATAAGATCGAGAAAGAATCAGCTCATG 547

Qy    542 TTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCG 601
            ||||| || || | ||| ||||| |||||||  || ||||| || || || ||   |
Db    548 CATGAATCAATCTCAGTGCTTCAGAAAAAGGATAGGGCATTGCAGGAGCAAAATAACCTG 607

Qy    602 CTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAG 647
          |||   || || || ||||||| ||||| |  |   |  |||||||
Db    608 CTTACAAAGAAAATAAAGGAGAAGGAGAGGGCACTAGCACAGCAAG 653


RESULT 12
AF130118
LOCUS       AF130118                956 bp    mRNA    linear   PLN 21-JUN-2001
DEFINITION  Capsicum annuum MADS box protein (MADS6) mRNA, complete cds.
ACCESSION   AF130118
VERSION     AF130118.2  GI:14518446
KEYWORDS    .
SOURCE      Capsicum annuum.
  ORGANISM  Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Capsicum.
REFERENCE   1  (bases 1 to 956)
  AUTHORS   Sung,S.K., Moon,Y.H., Chung,J.E., Lee,S.Y., Park,H.G. and An,G.
  TITLE     Characterization of MADS box genes from hot pepper
  JOURNAL   Mol. Cells 11 (3), 352-359 (2001)
  MEDLINE   21351847
   PUBMED   11459226
REFERENCE   2  (bases 1 to 956)
  AUTHORS   Jung,J.-Y., Moon,Y.-H., Chung,J.-E., Sung,S.-K. and An,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-FEB-1999) Department of Life Science, POSTECH, San
            31, Hyoja-dong, Nam-gu, Pohang 790-784, Republic of Korea
REFERENCE   3  (bases 1 to 956)
  AUTHORS   Jung,J.-Y., Moon,Y.-H., Chung,J.-E., Sung,S.-K. and An,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-JUN-2001) Department of Life Science, POSTECH, San
            31, Hyoja-dong, Nam-gu, Pohang 790-784, Republic of Korea
  REMARK    Sequence update by submitter
COMMENT     On Jun 21, 2001 this sequence version replaced gi:6651034.
FEATURES             Location/Qualifiers
     source          1. .956
                     /organism="Capsicum annuum"
                     /db_xref="taxon:4072"
```

```
                     /dev_stage="young flower buds"
     gene            1. .956
                     /gene="MADS6"
     CDS             55. .798
                     /gene="MADS6"
                     /function="transcription factor"
                     /note="member of the SQUAMOSA gene family"
                     /codon_start=1
                     /product="MADS box protein"
                     /protein_id="AAF22139.2"
                     /db_xref="GI:14518447"
                     /translation="MGRGRVQLRRIENKINRQVTFSKRRSGLLKKAHEISVLCDAEVG
                     LIVFSSKGKLFEYSTDSCMERILERYERYSYAERQLNATDVETPGSWTLEHAKLKARL
                     EVLQRNQRHYAGEDLDSLSMKELQNLEQQLDSALKHIRSRKNQLMHESISELQKKDKA
                     LQEQNNNLSKQMKEREKQLAQQHTPWEQQNHDHLNSSSFGLPHPFNNNHLGEVYPTAG
                     DNGEVEGSSRQQQQNAAVMPPWMLRHLNG"
BASE COUNT      333 a    173 c    213 g    237 t
ORIGIN

  Query Match             29.3%;  Score 311.6;  DB 8;  Length 956;
  Best Local Similarity   72.4%;  Pred. No. 7.7e-60;
  Matches  404;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy      88 AAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAA 147
           || |   ||| | |||||||||||| || |||||  ||| |||||| || || ||||| ||
Db      42 AACACTAAGAAAAATGGGAAGAGGAAGAGTTCAATTGAGGAGGATTGAAAATAAGATAAA 101

Qy     148 TAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTC 207
           ||||||||| ||||| || |||||  | |||||||| | ||||||||||||||||||||
Db     102 TAGGCAAGTGACTTTTTCGAAGAGGCGATCTGGTTTGTTGAAGAAAGCTCATGAGATCTC 161

Qy     208 TGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGA 267
           ||| || || |||||||| ||||   | || || || ||||| ||||| ||||| || ||
Db     162 TGTCCTTTGTGATGCTGAAGTTGGCTTGATTGTTTTTTCTTCTAAAGGGAAACTATTTGA 221

Qy     268 ATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTC 327
           ||||| || |||||||||||||| ||||| |||||  | |||||  | || | |||| |
Db     222 GTATTCTACTGACTCTTGCATGGAAAGGATTCTTGAGAGGTATGAAAGGTACTCATATGC 281

Qy     328 AGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGC 387
           || |  || |||  ||     || ||    |      |  | ||||  | | |||||||||
Db     282 TGAGAGGCAGCTTAATGCAACTGATGTCGAAACCCCGGGGAGTTGGACTTTGGAACATGC 341

Qy     388 TAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGA 447
           |||||| ||||| ||  |||||||  |  | |  ||| ||||| ||| |  ||| |||||
Db     342 TAAGCTTAAGGCCAGGCTTGAGGTTTTGCAAAGAAACCAAAGGCATTATGCGGGAGAAGA 401

Qy     448 TCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTAT 507
             | || || |||||  |||| ||||| || |   ||||||| || |||||| | ||| |
Db     402 CTTGGACTCATTGAGTATGAAAGAGCTTCAGAATCTGGAGCAGCAACTCGATTCTGCTCT 461

Qy     508 CAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAA 567
           ||| |||| | |||||||||||||||   |||   |||||||| |||| ||| || ||
Db     462 TAAGCACATTCGATCAAGAAAGAACCAATTGATGCATGAATCCATTTCTGAGCTGCAAAA 521

Qy     568 GAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGA 627
           |||||| || || |||||||| || |||||    |||   ||  |||| ||||| |||||
Db     522 GAAGGACAAGGCATTGCAAGAACAAAACAACAATCTTTCAAAGCAGATGAAGGAAAGGGA 581

Qy     628 GAAGAAAACGGGTCAGCA 645
           |||  |   ||  |||||
Db     582 GAAACAGCTGGCCCAGCA 599


RESULT 13
AF305696
LOCUS       AF305696               1032 bp    mRNA    linear   PLN 29-OCT-2000
DEFINITION  Eucalyptus globulus MADS-box protein EAP2S mRNA, complete cds,
            alternatively spliced.
ACCESSION   AF305696
VERSION     AF305696.1  GI:11037009
KEYWORDS    .
SOURCE      Eucalyptus globulus.
  ORGANISM  Eucalyptus globulus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Myrtales; Myrtaceae; Eucalyptus.
REFERENCE   1  (bases 1 to 1032)
  AUTHORS   Kyozuka,J., Harcourt,R., Peacock,W.J. and Dennis,E.S.
  TITLE     Eucalyptus has functional equivalents of the Arabidopsis AP1 gene
  JOURNAL   Plant Mol. Biol. 35 (5), 573-584 (1997)
  MEDLINE   98009992
   PUBMED   9349279
REFERENCE   2  (bases 1 to 1032)
  AUTHORS   Kyozuka,J., Harcourt,R., Peacock,W.J. and Dennis,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-SEP-2000) Plant Industry, CSIRO, GPO Box 1600,
            Canberra, ACT 2601, Australia
FEATURES             Location/Qualifiers
     source          1. .1032
                     /organism="Eucalyptus globulus"
                     /db_xref="taxon:34317"
     CDS             184. .801
                     /note="alternatively spliced"
                     /codon_start=1
                     /product="MADS-box protein EAP2S"
                     /protein_id="AAG27459.1"
                     /db_xref="GI:11037010"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRCGLLKKAHEISVLCDADVA
                     LIVFSTKGKLFEYATDCCMERILERYERYSYAESQVLTNNAETNGNWTLEHAKLKARM
                     EILQKNQKNLMGEELDSLSLKELQNLEHQLDTALKNIRSRKIQLMCESISELQRKDKA
                     LQEQNNMLAKKVKEKEKALAQQTQWDNPQDDGLTSSSGSLPSEWN"
BASE COUNT      302 a    229 c    264 g    237 t
ORIGIN

  Query Match             29.3%;  Score 311.2;  DB 8;  Length 1032;
  Best Local Similarity   73.0%;  Pred. No. 9.4e-60;
  Matches  400;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy      99 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 158
           | ||||| || || ||||| ||||||||||||||||||||||||||| || |||||||| |
Db     182 AAATGGGGAGGGGGAGGGTGCAGCTGAAGAGGATAGAGAACAAGATAAACAGGCAAGTGA 241

Qy     159 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 218
           | ||||| ||| | || |  ||  |||| ||||| || || |||||||| || |||||||
Db     242 CGTTCTCGAAGCGGAGATGCGGGCTGCTGAAGAAGGCCCACGAGATCTCCGTCCTCTGCG 301

Qy     219 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 278
           | || || || || ||||||||||||||  |||| ||||| |||||||| ||  ||||||
Db     302 ACGCCGACGTCGCCCTCATCGTCTTCTCCACCAAGGGCAAGCTCTTCGAGTACGCCACCG 361

Qy     279 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 338
           ||| ||||||||||||| || || || |||||  | |||| |||| |||| |  ||
Db     362 ACTGTTGCATGGAGAGGATCCTCGAGCGTTATGAGAGATATTCATATGCAGAGAGCCAGG 421

Qy     339 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 398
           || |       | |    | ||  || |||  | | |||||||| || |||||||
Db     422 TTCTCACAAACAATGCCGAAACCAATGGGAACTGGACTTTGGAACATGCAAAACTCAAGG 481

Qy     399 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 458
           | ||  | ||| | ||  |||| ||  ||| |||| | ||||| ||||| || |||||
Db     482 CCAGGATGGAGATTCTGCAGAAAAATCAAAAGAATCTGATGGGAGAAGAACTCGATTCTC 541

Qy     459 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 518
           | ||| | || ||||| || |  |||||||||||| ||  | ||| | |||| ||||
Db     542 TAAGCCTCAAAGAGCTTCAGAATTTGGAGCATCAGCTTGACACTGCTCTTAAGAACATTC 601

Qy     519 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 578
           |||| |||||||  |||   ||||  |||||||| |||| ||| ||||  |||||||| |
Db     602 GGTCTAGAAAGATTCAACTCATGTGTGAATCCATCTCTGAGCTTCAGAGAAAGGATAAGG 661

Qy     579 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 638
           |  ||||||| || |||||   |||    || ||| | || |||| |||||||  |  ||
Db     662 CGCTGCAAGAACAAAACAACATGCTGGCAAAGAAGGTGAAAGAGAAGGAGAAGGCACTGG 721

Qy     639 GTCAGCAA 646
```

```
Db    722 CACAGCAA 729


RESULT 14
AF306349
LOCUS       AF306349               1070 bp    mRNA    linear   PLN 08-NOV-2000
DEFINITION  Eucalyptus globulus MADS box protein AP2L mRNA, complete cds.
ACCESSION   AF306349
VERSION     AF306349.1  GI:11120556
KEYWORDS    .
SOURCE      Eucalyptus globulus.
  ORGANISM  Eucalyptus globulus
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Myrtales; Myrtaceae; Eucalyptus.
REFERENCE   1  (bases 1 to 1070)
  AUTHORS   Kyozuka,J., Harcourt,R., Peacock,W.J. and Dennis,E.S.
  TITLE     Eucalyptus has functional equivalents of the Arabidopsis AP1 gene
  JOURNAL   Plant Mol. Biol. 35 (5), 573-584 (1997)
  MEDLINE   98009992
   PUBMED   9349279
REFERENCE   2  (bases 1 to 1070)
  AUTHORS   Kyozuka,J., Harcourt,R., Peacock,W.J. and Dennis,E.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-SEP-2000) Plant Industry, CSIRO, GPO Box 1600,
            Canberra, ACT 2601, Australia
FEATURES             Location/Qualifiers
     source          1. .1070
                     /organism="Eucalyptus globulus"
                     /db_xref="taxon:34317"
     CDS             184. .921
                     /codon_start=1
                     /product="MADS box protein AP2L"
                     /protein_id="AAG30923.1"
                     /db_xref="GI:11120557"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRRCGLLKKAHEISVLCDADVA
                     LIVFSTKGKLFEYATDCCMERILERYERYSYAESQVLTNNAETNGNWTLEHAKLKARM
                     EILQKNQKNLMGEELDSLSLKELQNLEHQLDTALKNIRSRKIQLMCESISELQRKDKA
                     LQEQNNMLAKKVKEKEKALAQQTQWDNPQDDGLTSSSVILSQSLQPLNIGGPYHPSGI
                     EEGAALGPPQHRNATLFPSWMLSHLQE"
BASE COUNT      309 a    240 c    269 g    252 t
ORIGIN

  Query Match           29.3%;  Score 311.2;  DB 8;  Length 1070;
  Best Local Similarity   73.0%;  Pred. No. 9.4e-60;
  Matches  400;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy      99 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 158
           | ||||| || || ||||| ||||||||||||||||||||||||||||| || |||||||| |
Db     182 AAATGGGGAGGGGGAGGGTGCAGCTGAAGAGGATAGAGAACAAGATAAACAGGCAAGTGA 241

Qy     159 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 218
           | ||||| ||| | || |  ||  |||| ||||| || || |||||||| || |||||||
Db     242 CGTTCTCGAAGCGGAGATGCGGGCTGCTGAAGAAGGCCCACGAGATCTCCGTCCTCTGCG 301

Qy     219 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 278
           | || || || ||||||||||||||  |||| ||||| |||||||| ||  ||||||
Db     302 ACGCCGACGTCGCCCTCATCGTCTTCTCCACCAAGGGCAAGCTCTTCGAGTACGCCACCG 361

Qy     279 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 338
           ||| ||||||||||||||| || || || |||||  | |||| |||| |||| |  ||
Db     362 ACTGTTGCATGGAGAGGATCCTCGAGCGTTATGAGAGATATTCATATGCAGAGAGCCAGG 421

Qy     339 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 398
           || |       | |    |  | ||  || |||  | | |||||||| || |||||||
Db     422 TTCTCACAAACAATGCCGAAACCAATGGGAACTGGACTTTGGAACATGCAAAACTCAAGG 481

Qy     399 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 458
           | ||  | ||| | ||  |||| ||  ||| |||| | ||||| ||||| || ||||
Db     482 CCAGGATGGAGATTCTGCAGAAAAATCAAAAGAATCTGATGGGAGAAGAACTCGATTCTC 541

Qy     459 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 518
           | ||| | || ||||| || |  |||||||||||||| ||  | ||| | |||| ||||
Db     542 TAAGCCTCAAAGAGCTTCAGAATTTGGAGCATCAGCTTGACACTGCTCTTAAGAACATTC 601

Qy     519 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 578
           |||| |||||||  |||   ||||  |||||||| |||| ||| |||| |||||||| |
Db     602 GGTCTAGAAAGATTCAACTCATGTGTGAATCCATCTCTGAGCTTCAGAGAAAGGATAAGG 661

Qy     579 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 638
           |  ||||||| || |||||   |||    || ||| | || |||| |||||||  |  ||
Db     662 CGCTGCAAGAACAAAACAACATGCTGGCAAAGAAGGTGAAAGAGAAGGAGAAGGCACTGG 721

Qy     639 GTCAGCAA 646
           ||||||
Db     722 CACAGCAA 729


RESULT 15
AF068725
LOCUS       AF068725               1029 bp    mRNA    linear   PLN 16-JUN-1999
DEFINITION  Nicotiana sylvestris MADS-box protein MADS1 (NsMADS1) mRNA,
            complete cds.
ACCESSION   AF068725
VERSION     AF068725.1  GI:5070141
KEYWORDS    .
SOURCE      Nicotiana sylvestris.
  ORGANISM  Nicotiana sylvestris
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Nicotiana.
REFERENCE   1  (bases 1 to 1029)
  AUTHORS   Jang,S. and An,G.
  TITLE     NsMADS1, a member of the MADS gene family from Nicotiana sylvestris
  JOURNAL   J. Plant Biol. 42 (1), 85-87 (1999)
REFERENCE   2  (bases 1 to 1029)
  AUTHORS   Jang,S. and An,G.
  TITLE     Direct Submission
  JOURNAL   Submitted (28-MAY-1998) Department of Life Science, Pohang
            University of Science and Technology, San 31 Hyojadong, Pohang,
            Kyungbuk 790 - 784, Korea
FEATURES             Location/Qualifiers
     source          1. .1029
                     /organism="Nicotiana sylvestris"
                     /db_xref="taxon:4096"
                     /dev_stage="young floral bud"
     gene            1. .1029
                     /gene="NsMADS1"
     5'UTR           1. .79
                     /gene="NsMADS1"
     CDS             80. .817
                     /gene="NsMADS1"
                     /codon_start=1
                     /product="MADS-box protein MADS1"
                     /protein_id="AAD39036.1"
                     /db_xref="GI:5070142"
                     /translation="MGRGRVQLKRIENKINRQVTFSKRASGLLKKAHEISVLCDAEVG
                     LIVFSTKGKLFEYSTDSCMERILERYERYSYAERQLTATDDETPGSWTLEHAKLKARL
                     EVLQRNQRHYAGEDLDSLSMKELQNLEHQLDSALKHIRSRKNQLMHESISELQKKDKA
                     LQEQNNNLSKQVKEREKELAQQTQWEQQSHDHLNSSSFVLTQPLSSLHLGEAYPTAGD
                     NGEVEGSSRQQQQNTVMPPWMLRHLNG"
     3'UTR           818. .1029
                     /gene="NsMADS1"
     polyA_site      1020
                     /gene="NsMADS1"
BASE COUNT      340 a    189 c    216 g    284 t
ORIGIN

  Query Match           29.2%;  Score 310.2;  DB 8;  Length 1029;
  Best Local Similarity   69.2%;  Pred. No. 1.6e-59;
  Matches  423;  Conservative    0;  Mismatches  188;  Indels    0;  Gaps    0;

Qy      35 ATACTTTGGTCATTTCAGGGTTGTCGTTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAG 94
           ||   | | || | |   | ||      |  ||| | |   |||  | |||  | | |
Db      14 ATCACTCGATCTTGTGCAGCTTAGTACTCTTCTTTTTCTGTCTGAAAATTAAGTAAATTT 73
```

```
Qy      95 AGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAA 154
Db      74 GTAATAATGGGGAGAGGAAGAGTGCAACTGAAGAGAATTGAGAACAAGATCAATCGACAA 133

Qy     155 GTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTC 214
Db     134 GTCACCTTCTCAAAAAGAGCATCTGGTTTGCTTAAGAAAGCTCATGAAATCTCTGTGCTT 193

Qy     215 TGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCC 274
Db     194 TGTGATGCTGAGGTTGGTTTAATTGTTTTTTCTACTAAAGGGAAACTCTTTGAGTATTCC 253

Qy     275 ACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAA 334
Db     254 ACTGATTCTTGCATGGAAAGGATTCTTGAAAGGTATGAAAGGTACTCATATGCTGAGAGG 313

Qy     335 CAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTC 394
Db     314 CAGCTTACTGCTACTGATGATGAAACCCCGGGGAGCTGGACTTTGGAACATGCTAAGCTT 373

Qy     395 AAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGAT 454
Db     374 AAGGCCAGACTTGAGGTTTTGCAAAGAAACCAAAGGCATTATGCAGGAGAAGATTTGGAC 433

Qy     455 TCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGC 514
Db     434 TCATTAAGTATGAAAGAGCTTCAGAATCTTGAGCACCAGCTCGATTCTGCTCTTAAGCAC 493

Qy     515 ATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGAT 574
Db     494 ATTCGATCAAGAAAGAATCAATTGATGCATGAATCCATTTCTGAGCTGCAAAAGAAGGAC 553

Qy     575 AAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAA 634
Db     554 AAGGCATTGCAAGAGCAAAACAACAATCTCTCAAAGCAGGTGAAAGAAAGGGAGAAAGAG 613

Qy     635 ACGGGTCAGCA 645
Db     614 CTAGCTCAGCA 624
```

Search completed: February 2, 2003, 07:35:56
Job time : 3149 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: February 2, 2003, 00:08:41 ; Search time 272 Seconds
(without alignments)
8792.729 Million cell updates/sec

Title: US-09-981-087A-1
Perfect score: 1062
Sequence: 1 cccagagagacataagaaag.....taatgatttctcatgaaata 1062

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1062	100.0	1062	20	AAX18594	Arabidopsis AGL8 n
2	1062	100.0	1062	20	AAV99856	Arabidopsis AGL8-l
3	1055.8	99.4	1123	21	AAC41877	Arabidopsis thalia
4	444.8	41.9	542	21	AAC36764	Arabidopsis thalia
5	314.4	29.6	1032	17	AAT34430	Eucalyptus SQE2S c
6	312.4	29.4	1084	17	AAT34429	Eucalyptus SQE1 cD
7	312	29.4	1223	21	AAC55901	Eucalyptus grandis
8	311.2	29.3	1070	17	AAT34431	Eucalyptus SQE2L c
9	309.8	29.2	1131	20	AAX17796	Petunia MADS box t
10	305.8	28.8	738	22	AAC90635	Strawberry floweri
11	301	28.3	794	19	AAV58316	Brassica oleracea
12	301	28.3	794	19	AAT86629	APETALA1 gene from
13	301	28.3	794	19	AAV02761	Brassica oleracea
14	301	28.3	794	19	AAV06019	Brassica floral me
15	301	28.3	794	21	AAC61408	cDNA encoding a AP
16	301	28.3	794	21	AAZ57055	B. oleracea AP1 pr
17	301	28.3	794	21	AAZ92142	Brassica oleracea
18	297.8	28.0	794	18	AAT76886	Brassica oleracea
19	296.8	27.9	1057	19	AAT86628	APETALA1 gene from
20	296.8	27.9	1057	19	AAV02760	Arabidopsis floral
21	296.8	27.9	1057	19	AAV06018	Arabidopsis floral
22	296.8	27.9	1057	21	AAZ57054	A. thaliana AP1 pr
23	296.8	27.9	1057	21	AAZ92141	APETALA1 (AP1) nuc
24	296.8	27.9	1163	21	AAC51102	Arabidopsis thalia
25	296.8	27.9	1165	21	AAC37677	Arabidopsis thalia
26	296.8	27.9	1215	18	AAT76885	Arabidopsis floral
27	296.8	27.9	1215	19	AAV58306	Arabidopsis thalia
28	296.8	27.9	1215	21	AAC61407	cDNA encoding a AP
29	296.8	27.9	1220	21	AAA15016	cDNA encoding the
30	291.4	27.4	768	18	AAT76887	Cauliflower floral
31	291.4	27.4	768	19	AAV58317	Brassica oleracea
32	291.4	27.4	768	19	AAT86630	APETALA1 gene from
33	291.4	27.4	768	19	AAV02762	Cauliflowera flora
34	291.4	27.4	768	19	AAV06020	Cauliflower floral
35	291.4	27.4	768	21	AAC61409	cDNA encoding a AP
36	291.4	27.4	768	21	AAZ57056	B. oleracea var. b
37	291.4	27.4	768	21	AAZ92143	Cauliflower APETAL
38	276.2	26.0	1239	21	AAZ49161	MADS box gene. Ge
39	271.6	25.6	966	22	AAH87721	Peppermint plant o
40	266	25.0	777	18	AAT76888	Arabidopsis floral
41	266	25.0	779	19	AAV58300	Arabidopsis thalia
42	266	25.0	779	19	AAT86632	CAULIFLOWER gene f
43	266	25.0	779	19	AAV06022	Arabidopsis floral
44	266	25.0	779	19	AAV02764	Arabidopsis floral
45	266	25.0	779	21	AAC61411	cDNA encoding a CA

ALIGNMENTS

RESULT 1
AAX18594
ID AAX18594 standard; DNA; 1062 BP.
XX
AC AAX18594;
XX
DT 06-MAY-1999 (first entry)
XX
DE Arabidopsis AGL8 nucleotide sequence.
XX
KW Arabidopsis; AGL8; agamous-like 8; seed plant; fruit size; ss.
XX
OS Arabidopsis sp.
XX
FH Key Location/Qualifiers
FT CDS 101..829
FT /*tag= a
XX
PN WO9900503-A1.
XX
PD 07-JAN-1999.
XX
PF 26-JUN-1998; 98WO-US13249.
XX
PR 27-JUN-1997; 97US-0051030.
XX
PA (COLD-) COLD SPRING HARBOR LAB.
PA (REGC) UNIV CALIFORNIA.
XX
PI Ferrandiz C, Gu Q, Martienssen R, Yanofsky MF;
XX
DR WPI; 1999-095748/08.

```
DR   P-PSDB; AAW98921.
XX
PT   Use of agamous-like 8 related nucleic acids - useful for producing
PT   transgenic seed plants which can produce seed or fruit of increased
PT   or decreased size
XX
PS   Disclosure; Fig 1; 68pp; English.
XX
CC   The present invention describes a non-naturally occurring seed plant
CC   comprising an ectopically expressed nucleic acid molecule encoding an
CC   agamous-like 8 (AGL8)-related gene product. The seed plant characterised
CC   by producing seeds of increased size. The present sequence encodes
CC   Arabidopsis AGL-8. The seed plant can be used for producing seed or
CC   fruit of increased or decreased size. The present invention can relate
CC   to seed plants such as, e.g. canola, soybean, pea, broad bean, peanut,
CC   coconut, sunflower, cotton, coffee, cacao, citrus, grape, apples,
CC   strawberry, olive and tomato.
XX
SQ   Sequence 1062 BP; 327 A; 193 C; 240 G; 302 T; 0 other;

  Query Match            100.0%;  Score 1062;  DB 20;  Length 1062;
  Best Local Similarity  100.0%;  Pred. No. 2.4e-262;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
           |||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 2
AAV99856
ID   AAV99856 standard; cDNA; 1062 BP.
XX
AC   AAV99856;
XX
DT   10-MAY-1999  (first entry)
XX
DE   Arabidopsis AGL8-like cDNA clone.
XX
KW   AGL8-like gene; agamous-like 8 gene; transcription factor;
KW   seed dispersal; dehiscence; transgenic plant; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key             Location/Qualifiers
FT   CDS             101..829
FT                   /*tag= a
FT   polyA_site      1062
FT                   /*tag= b
XX
PN   WO9900502-A1.
XX
PD   07-JAN-1999.
XX
PF   25-JUN-1998;   98WO-US13208.
XX
PR   28-APR-1998;   98US-0067800.
PR   27-JUN-1997;   97US-0051030.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Ferrandiz C,  Yanofsky MF;
XX
DR   WPI; 1999-095747/08.
DR   P-PSDB; AAW80998.
XX
PT   Use of agamous-like nucleic acids - useful for the production of
PT   transgenic seed plants in which dehiscence is modified resulting in
PT   delayed seed dispersal
XX
PS   Disclosure; Fig 6; 126pp; English.
XX
CC   This cDNA clone encodes an Arabidopsis agamous-like 8 (AGL8)-like
CC   gene product (see AAW80998) that regulates dehiscence.  The invention
CC   provides a transgenic seed plant that is characterised by delayed
CC   seed dispersal due to ectopic expression of a nucleic acid molecule
CC   encoding an AGL8-like gene product.  The ectopically expressed
CC   nucleic acid molecule may be operatively linked to an exogenous
```

```
CC   regulatory element such as a dehiscence zone-selective regulatory
CC   element derived from the Arabidopsis AGL1 gene (see AAV99857) or
CC   AGL5 gene (see AAV99858).  The transgenic seed plants include members
CC   of the Brassicaceae, such as rapeseed, and members of the Fabaceae,
CC   such as soybean, pea, lentil and bean.  A plant expression vector
CC   comprising a dehiscence zone-selective regulatory element, and a
CC   kit for producing the transgenic seed plants are also provided.
XX
SQ   Sequence 1062 BP; 327 A; 193 C; 240 G; 302 T; 0 other;

  Query Match             100.0%;  Score 1062;  DB 20;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 2.4e-262;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTTTGTGGAGAGAGTTGGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
          |||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 3
AAC41877
ID   AAC41877 standard; DNA; 1123 BP.
XX
AC   AAC41877;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 33470.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
```

PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 99.4%; Score 1055.8; DB 21; Length 1123;
Best Local Similarity 99.8%; Pred. No. 9.5e-261;

```
  Matches 1057;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      23 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 82

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      83 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 142

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     143 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 202

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     203 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 262

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 322

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     323 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 382

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     383 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 442

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| |||||
Db     443 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTTCAAAG 502

Qy     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     503 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 562

Qy     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     563 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 622

Qy     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     623 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 682

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     683 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 742

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     743 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 802

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     803 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 862

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     863 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 922

Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
          |||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||
Db     923 TTTGGTGACTTATACTCATTACTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 982

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     983 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1042

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAA 1059
          |||||||||||||||||||||||||||||||||||||||
Db    1043 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAA 1081
```

```
RESULT 4
AAC36764
ID   AAC36764 standard; DNA; 542 BP.
XX
AC   AAC36764;
XX
DT   17-OCT-2000  (first entry)
XX
DE   Arabidopsis thaliana DNA fragment SEQ ID NO: 15001.
XX
KW   Hybridisation assay; genetic mapping; gene expression control;
KW   protein identification; signal transduction pathway;
KW   metabolic pathway; promoter; termination sequence; ss.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
PF   25-FEB-2000; 2000EP-0301439.
XX
PR   25-FEB-1999;   99US-0121825.
PR   05-MAR-1999;   99US-0123180.
PR   09-MAR-1999;   99US-0123548.
PR   23-MAR-1999;   99US-0125788.
PR   25-MAR-1999;   99US-0126264.
PR   29-MAR-1999;   99US-0126785.
PR   01-APR-1999;   99US-0127462.
PR   06-APR-1999;   99US-0128234.
PR   08-APR-1999;   99US-0128714.
PR   16-APR-1999;   99US-0129845.
PR   19-APR-1999;   99US-0130077.
PR   21-APR-1999;   99US-0130449.
PR   23-APR-1999;   99US-0130510.
PR   23-APR-1999;   99US-0130891.
PR   28-APR-1999;   99US-0131449.
PR   30-APR-1999;   99US-0132048.
PR   30-APR-1999;   99US-0132407.
PR   04-MAY-1999;   99US-0132484.
PR   05-MAY-1999;   99US-0132485.
PR   06-MAY-1999;   99US-0132486.
PR   06-MAY-1999;   99US-0132487.
PR   07-MAY-1999;   99US-0132863.
PR   11-MAY-1999;   99US-0134256.
PR   14-MAY-1999;   99US-0134218.
PR   14-MAY-1999;   99US-0134219.
PR   14-MAY-1999;   99US-0134221.
PR   14-MAY-1999;   99US-0134370.
PR   18-MAY-1999;   99US-0134768.
PR   19-MAY-1999;   99US-0134941.
PR   20-MAY-1999;   99US-0135124.
PR   21-MAY-1999;   99US-0135353.
PR   24-MAY-1999;   99US-0135629.
PR   25-MAY-1999;   99US-0136021.
PR   27-MAY-1999;   99US-0136392.
PR   28-MAY-1999;   99US-0136782.
PR   01-JUN-1999;   99US-0137222.
PR   03-JUN-1999;   99US-0137528.
PR   04-JUN-1999;   99US-0137502.
PR   07-JUN-1999;   99US-0137724.
PR   08-JUN-1999;   99US-0138094.
PR   10-JUN-1999;   99US-0138540.
PR   10-JUN-1999;   99US-0138847.
PR   14-JUN-1999;   99US-0139119.
PR   16-JUN-1999;   99US-0139452.
PR   16-JUN-1999;   99US-0139453.
PR   17-JUN-1999;   99US-0139492.
PR   18-JUN-1999;   99US-0139454.
PR   18-JUN-1999;   99US-0139455.
PR   18-JUN-1999;   99US-0139456.
```

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 41.9%; Score 444.8; DB 21; Length 542;
Best Local Similarity 99.6%; Pred. No. 3.3e-104;
Matches 446; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 612 AGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCA 671
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 1 AGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCA 60

Qy 672 ACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGATGGCTTTGTGG 731
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 61 ACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGATGGCTTTGTGG 120

Qy 732 AGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTCTCTGCTTC 791
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

```
Db    121 AGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTCTCTGCTTC 180

Qy    792 CGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCACTCTTTATAATA 851
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    181 CGGCTTGGATGTTACGTCCTACCACTACGAACAAGTAGAACTATCTCACTCTTTATAATA 240

Qy    852 TAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTTTTTGGTGACTT 911
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 TAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTTTTTGGTGACTT 300

Qy    912 ATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTATGATGGAACTCC 971
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 ATACTCATTACTAATACCGATATGTTTTAGCTAGTCATATTATATGTATGATGGAACTCC 360

Qy    972 GTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTAAGAACAAAGAT 1031
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTAAGAACAAAGAT 420

Qy   1032 TCATATCTTGGTAATGATTTCTCATGAA 1059
          ||||||||||||||||||||||||||||
Db    421 TCATATCTTGGTAATGATTTCTCATGAA 448
```

RESULT 5
AAT34430
ID AAT34430 standard; cDNA; 1032 BP.
XX
AC AAT34430;
XX
DT 01-OCT-1996 (first entry)
XX
DE Eucalyptus SQE2S cDNA.
XX
KW Eucalyptus; reproductive gene; SQE2S; Arabidopsis; agomous gene;
KW Antirrhinum; plena gene; sterility; fertility; forest tree;
KW transgenic plant; flower development; antisense; ribozyme; ss.
XX
OS Eucalyptus globulus.
XX
FH Key Location/Qualifiers
FT CDS 184..801
FT /*tag= a
FT /product= SQE-2S
XX
PN AU9539013-A.
XX
PD 30-MAY-1996.
XX
PF 22-NOV-1995; 95AU-0039013.
XX
PR 22-NOV-1994; 94AU-0009589.
XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX
PI Dennis ES, Harcourt RL, Kyozuka J, Llewellyn D;
PI Peacock WJ, Southerton S;
XX
DR WPI; 1996-278411/29.
DR P-PSDB; AAR99635.
XX
PT Eucalyptus reproductive genes - useful for prodn. of sterile
PT Eucalyptus trees useful for establishing wood lot plantations or in
PT re-forestation projects
XX
PS Claim 8; Page 34-35; 60pp; English.
XX
CC cDNA clones SQE1 (AAT34429), SQE2S (AAT34430) and SQE2L (AAT34431) were
CC identified by homology to the MADS box of the Arabidopsis apetala
CC 1 (AP1) and Antirrhinum squamosa (SQUA) flower development genes.
CC They were obtd. by PCR amplification of E. globulus DNA using
CC primers based on the MADS box of the Arabidopsis and Antirrhinum
CC genes, followed by screening of a E. globulus young flower bud
CC library. The SQE2S cDNA codes for protein SQE2-S (AAR99635).
CC Antisense or ribozyme constructs of SQE, or of FLE2 and SQE genes
CC (see also AAT34426, AAT34428 and AAT34432), may be used to produce
CC sterile transgenic Eucalyptus trees by minimising inflorescence.
XX
SQ Sequence 1032 BP; 299 A; 231 C; 263 G; 239 T; 0 other;

```
  Query Match            29.6%;  Score 314.4;  DB 17;  Length 1032;
  Best Local Similarity  73.4%;  Pred. No. 1.1e-70;
  Matches  402;  Conservative    0;  Mismatches  146;  Indels    0;  Gaps    0;

Qy     99 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 158
          | ||||| || || ||||| ||||||||||||||||||||||||| || |||||||| |
Db    182 AAATGGGGAGGGGGAGGGTGCAGCTGAAGAGGATAGAGAACAAGATAAACAGGCAAGTGA 241

Qy    159 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 218
          | ||||| ||| | || |  ||  |||| ||||| || || |||||||| || |||||||
Db    242 CGTTCTCGAAGCGGAGATGCGGGCTGCTGAAGAAGGCCCACGAGATCTCCGTCCTCTGCG 301

Qy    219 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 278
          | || || || || ||||||||||||||  |||| ||||| |||||||| ||  ||||||
Db    302 ACGCCGACGTCGCCCTCATCGTCTTCTCCACCAAGGGCAAGCTCTTCGAGTACGCCACCG 361

Qy    279 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 338
          ||| |||||||||||||| || || || |||||| | |||| |||| |||| || ||
Db    362 ACTGTTGCATGGAGAGGATCCTCGAGCGTTATGATAGATATTCATATGCAGAGAACCAGG 421

Qy    339 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 398
          || |        | |    |   | ||  || |||  | | |||||||| || |||||||
Db    422 TTCTCACAAACAATGCCGAAACCAATGGGAACTGGACTTTGGAACATGCAAAACTCAAGG 481

Qy    399 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 458
          | ||  | ||| | ||  |||| ||  ||| |||| | ||||| ||||| || |||||
Db    482 CCAGGATGGAGATTCTGCAGAAAAATCAAAAGAATCTGATGGGAGAAGAACTCGATTCTC 541

Qy    459 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 518
          | ||| | || ||||| || |  |||||||||||||| ||  | ||| | |||| ||||
Db    542 TAAGCCTCAAAGAGCTTCAGAATTTGGAGCATCAGCTTGACACTGCTCTTAAGAACATTC 601

Qy    519 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 578
          |||| |||||||  |||   |||| ||||||||  |||| ||| |||| |||||||| |
Db    602 GGTCTAGAAAGATTCAACTCATGTGTGAATCCATCTCTGAGCTTCAGAGAAAGGATAAGG 661

Qy    579 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 638
          |  |||||| || |||||   |||    || ||| | || |||| |||||||  |  ||
Db    662 CGCTGCAAGAACAAAACAACATGCTGGCAAAGAAGGTGAAAGAGAAGGAGAAGGCACTGG 721

Qy    639 GTCAGCAA 646
            ||||||
Db    722 CACAGCAA 729
```

RESULT 6
AAT34429
ID AAT34429 standard; cDNA; 1084 BP.
XX
AC AAT34429;
XX
DT 01-OCT-1996 (first entry)
XX
DE Eucalyptus SQE1 cDNA.
XX
KW Eucalyptus; reproductive gene; SQE1; Arabidopsis; agomous gene;
KW Antirrhinum; plena gene; sterility; fertility; forest tree;
KW transgenic plant; flower development; antisense; ribozyme; ss.
XX
OS Eucalyptus globulus.
XX
FH Key Location/Qualifiers
FT CDS 107..841
FT /*tag= a
FT /product= SQE-1
XX
PN AU9539013-A.

XX
PD 30-MAY-1996.
XX
PF 22-NOV-1995; 95AU-0039013.
XX
PR 22-NOV-1994; 94AU-0009589.
XX
PA (CSIR) COMMONWEALTH SCI & IND RES ORG.
XX
PI Dennis ES, Harcourt RL, Kyozuka J, Llewellyn D;
PI Peacock WJ, Southerton S;
XX
DR WPI; 1996-278411/29.
DR P-PSDB; AAR99634.
XX
PT Eucalyptus reproductive genes - useful for prodn. of sterile
PT Eucalyptus trees useful for establishing wood lot plantations or in
PT re-forestation projects
XX
PS Claim 8; Page 31-32; 60pp; English.
XX
CC cDNA clones SQE1 (AAT34429), SQE2S (AAT34430) and SQE2L (AAT34431) were
CC identified by homology to the MADS box of the Arabidopsis apetala
CC 1 (AP1) and Antirrhinum squamosa (SQUA) flower development genes.
CC They were obtd. by PCR amplification of E. globulus DNA using
CC primers based on the MADS box of the Arabidopsis and Antirrhinum
CC genes, followed by screening of a E. globulus young flower bud
CC library. The SQE1 gene contains a MADS box with 95% homology
CC to AP1 and 91% homology to SQUA. Its predicted amino acid
CC sequence is given in AAR99634. Antisense or ribozyme constructs of
CC SQE, or of FLE2 and SQE genes (see also AAT34426, AAT34428 and
CC AAT34432), may be used to produce sterile transgenic Eucalyptus trees by
CC minimising inflorescence.
XX
SQ Sequence 1084 BP; 297 A; 258 C; 251 G; 278 T; 0 other;

Query Match 29.4%; Score 312.4; DB 17; Length 1084;
Best Local Similarity 70.8%; Pred. No. 3.8e-70;
Matches 415; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

```
Qy      62 TTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAG 121
           || |  |   |||  | |    | ||||  |        |||||| || || ||| | |||
Db      68 TTTTGGCGTGTGTGTTGGGTTGTGTGAATTGGAGCGAGGATGGGGAGGGGGAGGCTGCAG 127

Qy     122 CTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGT 181
           ||||||||||||||||||||||||||  | ||||| || ||||| ||||| ||| | |||
Db     128 CTGAAGAGGATAGAGAACAAGATCAACCGACAAGTCACCTTCTCCAAGAGGAGGGCGGGT 187

Qy     182 TTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTC 241
           |||||||||| || || |||||||| || |||||||| || ||||| || |||||||||
Db     188 CTGCTCAAGAAGGCCCACGAGATCTCCGTCCTCTGCGACGCCGAGGTCGCCCTCATCGTC 247

Qy     242 TTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACTT 301
           |||||  |||| ||||| |||||||| || |||||||| || |||||||||| || ||
Db     248 TTCTCCGCCAAGGGCAAGCTCTTCGAGTACTCCACCGATTCCTGCATGGAGAGAATTCTC 307

Qy     302 GAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAA 361
           ||||||||||  | || | |||| | ||  | ||| || |||   | ||      |
Db     308 GAACGCTATGAAAGATACTCATATGCGGAGCACCAAGTTCTTGCAAGCGAGACGGAATCG 367

Qy     362 AGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAG 421
           | ||  |  |||  | | || ||||||||||||||||| ||| |||| ||  |  | |
Db     368 ATTGGTAGCTGGACTTTGGAGCATGCTAAGCTCAAGGCCAGACTTGAAGTTTTACACAGA 427

Qy     422 AACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGC 481
           ||  | ||| |||| ||||| |||||||||||||| |||||  | |||||  |||||||
Db     428 AATTATAGGCATTTCATGGGCGAAGATCTTGATTCTTTGAGTCTCAAGGACCTCCAAAAT 487

Qy     482 TTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATG 541
           |||||||| || || ||  | ||| | ||   ||| || || |||||||| ||    |||
Db     488 TTGGAGCAGCAACTGGAGTCTGCTCTTAAACACATAAGATCGAGAAAGAATCAGCTCATG 547

Qy     542 TTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCG 601
```

```
           ||||| || || | | | ||||| |||||||  || ||||| || || || ||   |
Db     548 CATGAATCAATCTCAGTGATTCAGAAAAAGGATAGGGCATTGCAGGAGCAAAATAACCTG 607

Qy     602 CTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAG 647
           |||   || || || ||||||| ||||| |  |   |  |||||||
Db     608 CTTACAAAGAAAATAAAGGAGAAGGAGAGGGCACTAGCACAGCAAG 653
```

RESULT 7
AAC55901
ID AAC55901 standard; DNA; 1223 BP.
XX
AC AAC55901;
XX
DT 25-JAN-2001 (first entry)
XX
DE Eucalyptus grandis transcription factor DNA sequence #32.
XX
KW Plant; transcription factor; gene expression; eucalyptus; pine; acacia;
KW poplar; sweetgum; teak; mahogany; bZIP; G-box binding factor;
KW basic helix-loop-helix zipper; homeotic; homeodomain; homeobox; MADS;
KW homeodomain zipper; LIM domain; AP2; EREBs; zinc finger domain;
KW type 2 Cys2His2; CCAAT box element; MYB; ss.
XX
OS Eucalyptus grandis.
XX
PN WO200053724-A2.
XX
PD 14-SEP-2000.
XX
PF 09-MAR-2000; 2000WO-US06112.
XX
PR 11-MAR-1999; 99US-0266513.
PR 18-AUG-1999; 99US-0149485.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Wood M, McGrath A, Shenk MA, Glenn M;
XX
DR WPI; 2000-579369/54.
XX
PT New isolated polynucleotide encoding a plant transcription factor for
PT producing a plant e.g. a woody plant, preferably eucalyptus or pine,
PT having modified gene expression or modified activity of a polypeptide
PT -
XX
PS Claim 1; Pages 51-52; 747pp; English.
XX
CC The present invention relates to novel plant transcription factors from
CC Eucalyptus grandis or Pinus radiata. The present sequence is the coding
CC sequence for one such transcription factor. The transcription factor may
CC be used to produce a plant having modified gene expression such as a
CC woody plant e.g. a eucalyptus, pine, acacia, poplar, sweetgum, teak, or
CC mahogany species or to modify the activity of a polypeptide in a plant.
CC The transcription factors of the present invention are members from the
CC following families of regulatory proteins: bZIP, bZIP family of G-box
CC binding factors, basic helix-loop-helix zipper,
CC homeotic/homeodomain/homeobox/MADS, homeodomain zipper, LIM domain, AP2
CC and EREBs, zinc finger domains of type 2 Cys2His2, CCAAT box elements
CC and MYB.
XX
SQ Sequence 1223 BP; 334 A; 292 C; 283 G; 314 T; 0 other;

Query Match 29.4%; Score 312; DB 21; Length 1223;
Best Local Similarity 71.4%; Pred. No. 5e-70;
Matches 411; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

```
Qy      72 TGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGA 131
           |||  | |    | ||||  |        ||||| || || ||  | ||||||||||||
Db     170 TGTGTTGGGTTGTGTGAATTGGAGCGAGGATGGGGAGGGGGAGACTGCAGCTGAAGAGGA 229

Qy     132 TAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGA 191
```

```
Db    230 TAGAGAACAAGATCAACCGGCAAGTCACCTTCTCCAAGAGGAGGGCGGGTCTGCTCAAGA 289

Qy    192 AAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCA 251

Db    290 AGGCCCACGAGATCTCCGTACTCTGCGACGCCGAGGTCGCCCTCATCATCTTCTCCGCCA 349

Qy    252 AAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATG 311

Db    350 AGGGCAAGCTCTTCGAGTACTCCACCGATTCCTGCATGGAGAGAATTCTCGAACGCTATG 409

Qy    312 ATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATT 371

Db    410 AAAGATACTCATATTCGGAGCACCAAGTTCTTGCAAGTGAGACAGAATCGATTGGTAGCT 469

Qy    372 GGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGA 431

Db    470 GGACTTTGGAGCATGCTAAGCTCAAGGCCAGACTTGAAGTTTTACACAGAAATTATAGGC 529

Qy    432 ATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATC 491

Db    530 ATTTCATGGGAGAAGATCTTGATTCTTTGAGTCTCAAGGACCTCCAAAATTTGGAGCAGC 589

Qy    492 AGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCA 551

Db    590 AACTGGAGTCTGCTCTTAAACACATAAGATCGAGAAAGAATCAGCTCATGCATGAATCAA 649

Qy    552 TATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAA 611

Db    650 TCTCAGTGCTTCAGAAAAAGGATAGGGCATTGCAGGAGCAAAACAACCTGCTTACAAGGA 709

Qy    612 AGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAG 647

Db    710 AAGTAAAGGAGAAGGAGAGGGCACTAGCGCAGCAAG 745
```

```
RESULT 8
AAT34431
ID   AAT34431 standard; cDNA; 1070 BP.
XX
AC   AAT34431;
XX
DT   01-OCT-1996  (first entry)
XX
DE   Eucalyptus SQE2L cDNA.
XX
KW   Eucalyptus; reproductive gene; SQE2L; Arabidopsis; agomous gene;
KW   Antirrhinum; plena gene; sterility; fertility; forest tree;
KW   transgenic plant; flower development; antisense; ribozyme; ss.
XX
OS   Eucalyptus globulus.
XX
FH   Key             Location/Qualifiers
FT   CDS             184..921
FT                   /*tag= a
FT                   /product= SQE-2L
XX
PN   AU9539013-A.
XX
PD   30-MAY-1996.
XX
PF   22-NOV-1995;   95AU-0039013.
XX
PR   22-NOV-1994;   94AU-0009589.
XX
PA   (CSIR ) COMMONWEALTH SCI & IND RES ORG.
XX
PI   Dennis ES,  Harcourt RL,  Kyozuka J,  Llewellyn D;
PI   Peacock WJ,  Southerton S;
XX
DR   WPI; 1996-278411/29.
DR   P-PSDB; AAR99636.
XX
PT   Eucalyptus reproductive genes - useful for prodn. of sterile
PT   Eucalyptus trees useful for establishing wood lot plantations or in
PT   re-forestation projects
XX
PS   Claim 8; Page 36-37; 60pp; English.
XX
CC   cDNA clones SQE1 (AAT34429), SQE2S (AAT34430) and SQE2L (AAT34431) were
CC   identified by homology to the MADS box of the Arabidopsis apetala
CC   1 (AP1) and Antirrhinum squamosa (SQUA) flower development genes.
CC   They were obtd. by PCR amplification of E. globulus DNA using
CC   primers based on the MADS box of the Arabidopsis and Antirrhinum
CC   genes, followed by screening of a E. globulus young flower bud
CC   library.  The SQE2L cDNA codes for protein SQE2-L (AAR99636).
CC   Antisense or ribozyme constructs of SQE, or of FLE2 and SQE genes
CC   (see also AAT34426, AAT34428 and AAT34432), may be used to produce
CC   sterile transgenic Eucalyptus trees by minimising inflorescence.
XX
SQ   Sequence 1070 BP; 308 A; 240 C; 270 G; 252 T; 0 other;

  Query Match             29.3%;  Score 311.2;  DB 17;  Length 1070;
  Best Local Similarity   73.0%;  Pred. No. 7.6e-70;
  Matches  400;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy     99 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 158

Db    182 AAATGGGGAGGGGGAGGGTGCAGCTGAAGAGGATAGAGAACAAGATAAACAGGCAAGTGA 241

Qy    159 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 218

Db    242 CGTTCTCGAAGCGGAGATGCGGGCTGCTGAAGAAGGCCCACGAGATCTCCGTCCTCTGCG 301

Qy    219 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 278

Db    302 ACGCCGACGTCGCCCTCATCGTCTTCTCCACCAAGGGCAAGCTCTTCGAGTACGCCACCG 361

Qy    279 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 338

Db    362 ACTGTTGCATGGAGAGGATCCTCGAGCGTTATGAGAGATATTCATATGCAGAGAGCCAGG 421

Qy    339 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 398

Db    422 TTCTCACAAACAATGCCGAAACCAATGGGAACTGGACTTTGGAACATGCAAAACTCAAGG 481

Qy    399 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 458

Db    482 CCAGGATGGAGATTCTGCAGAAAAATCAAAAGAATCTGATGGGAGAAGAACTCGATTCTC 541

Qy    459 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 518

Db    542 TAAGCCTCAAAGAGCTTCAGAATTTGGAGCATCAGCTTGACACTGCTCTTAAGAACATTC 601

Qy    519 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 578

Db    602 GGTCTAGAAAGATTCAACTCATGTGTGAATCCATCTCTGAGCTTCAGAGAAAGGATAAGG 661

Qy    579 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 638

Db    662 CGCTGCAAGAACAAAACAACATGCTGGCAAAGAAGGTGAAAGAGAAGGAGAAGGCACTGG 721

Qy    639 GTCAGCAA 646

Db    722 CACAGCAA 729
```

```
RESULT 9
AAX17796
ID   AAX17796 standard; cDNA to mRNA; 1131 BP.
XX
AC   AAX17796;
XX
DT   21-MAY-1999  (first entry)
XX
DE   Petunia MADS box transcription factor, FBP10 encoding cDNA.
```

```
XX
KW   Petunia; MADS box; transcription factor; FBP10; flowering; inhibition;
KW   inflorescence; vegetative organ; storage organ; wood production;
KW   nitrogen fixation; ornamental plant; ss.
XX
OS   Petunia sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             65..805
FT                   /*tag=  a
FT                   /product= "FBP10 protein"
FT   misc_feature    65..232
FT                   /*tag=  b
FT                   /note= "sequence encoding MADS box region"
XX
PN   WO9904003-A1.
XX
PD   28-JAN-1999.
XX
PF   18-JUL-1997;   97WO-NL00424.
XX
PR   18-JUL-1997;   97WO-NL00424.
XX
PA   (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
XX
PI   Angenent GC,  Busscher M,  Franken J;
XX
DR   WPI; 1999-132247/11.
DR   P-PSDB; AAW95018.
XX
PT   New DNA sequence encoding a MADS box transcription factor, FBP10 -
PT   useful for producing transgenic plants in which flowering is
PT   inhibited
XX
PS   Claim 2; Fig 1; 38pp; English.
XX
CC   This cDNA encodes a petunia MADS box transcription factor, FBP10 protein.
CC   This can be used in a process of producing a transgenic plant in which
CC   flowering is inhibited due to the expression of endogenous FBP10 or a
CC   homologous gene being inhibited; This is useful for producing plants
CC   without inflorescences or flowers which is beneficial for directing the
CC   high energy, which seeds and flowers utilise, to form vegetative organs
CC   or deposition in storage organs. Therefore, the total biomass of
CC   vegetative tissues such as roots, tubers, stem and leaves can be
CC   increased. This is particularly useful for plants used economically e.g.
CC   lettuce, spinach, chicory, sugar beet, potato, trees for wood production,
CC   tobacco, grasses, plants for nitrogen fixation, ornamental plants for
CC   production of cuttings.
XX
SQ   Sequence 1131 BP; 383 A; 185 C; 229 G; 334 T; 0 other;

  Query Match             29.2%;  Score 309.8;  DB 20;  Length 1131;
  Best Local Similarity   70.6%;  Pred. No. 1.8e-69;
  Matches  413;  Conservative    0;  Mismatches  172;  Indels    0;  Gaps    0;

Qy      62 TTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAG 121
           ||   | | || | |   | | |||   |  ||| | | ||||||||||| || || |||
Db      26 TTTAATTTATTCTGCACTATACTTTTTTGTTAGACAAAAATGGGAAGAGGAAGAGTGCAG 85

Qy     122 CTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGT 181
           ||||||| || ||||| || || ||||| |||||||||| |||||  |  | |||||
Db      86 ATGAAGAGAATTGAGAATAAAATTAATAGACAAGTTACTTTTTCAAAACGTCGATCTGGA 145

Qy     182 TTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTC 241
           ||  | |||||||||||||| |||||||| || || |||||||| |||| | | || ||
Db     146 TTATTGAAGAAAGCTCATGAAATCTCTGTGCTTTGTGATGCTGAAGTTGGTTTAATTGTT 205

Qy     242 TTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACTT 301
           || ||| | |||||||||||||| || ||| | || || ||||||||||||||||| |||
Db     206 TTTTCTACTAAAGGCAAACTCTTTGAGTATGCTACTGATTCTTGCATGGAGAGGATTCTT 265

Qy     302 GAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAA 361
           ||| | |||||  | || | |||| | || |  || ||||||     ||   |
.           . .
```

```
Db     266 GAAAGATATGAAAGATACTCATATGCTGAGAGGCAGCTTGTTTCTACTGATCATAGCTCC 325

Qy     362 AGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAG 421
           | ||  |||  ||| |||||||| || || ||||| ||| |||||||  |  |||
Db     326 CCGGGAAGCTGGAATCTGGAACATGCAAAACTTAAGGCCAGAATTGAGGTTGTGCAGAGA 385

Qy     422 AACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGC 481
           ||| ||||| ||| |||||| |||||| | || ||||| ||  |||| || || || |
Db     386 AACCAAAGGCATTATATGGGAGAAGATTTGGACTCGTTAAGTATGAAAGACCTTCAGAAT 445

Qy     482 TTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATG 541
           || || || ||||| ||| |  || | ||   |||| | ||||||||||||||||   |||
Db     446 TTAGAACAACAGCTGGATTCTTCTCTTAAACACATTCGATCAAGAAAGAACCAATTGATG 505

Qy     542 TTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCG 601
             || ||||| |||| ||| || || ||||| ||| | ||||||| || |||||
Db     506 CATGAGTCCATTTCTGAGCTTCAAAAAAAGGACAAATCATTGCAAGAGCAAAACAACCTT 565

Qy     602 CTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAA 646
           |||   || ||| | ||||||||||||||  |   || ||| |||
Db     566 CTTTCAAAGAAGGTGAAGGAGAGGGAGAAAGAGTTGGCTCAACAA 610


RESULT 10
AAC90635
ID   AAC90635 standard; cDNA; 738 BP.
XX
AC   AAC90635;
XX
DT   20-MAR-2001  (first entry)
XX
DE   Strawberry flowering regulation protein coding sequence SEQ ID NO: 3.
XX
KW   Strawberry; flowering regulation; floral homeotic gene; field crop;
KW   harvesting; fruit production; ss.
XX
OS   Fragaria vesca.
XX
PN   WO200071722-A1.
XX
PD   30-NOV-2000.
XX
PF   24-MAY-2000; 2000WO-US14297.
XX
PR   25-MAY-1999;   99US-0318789.
PR   24-MAY-2000; 2000US-0318789.
XX
PA   (DNAP ) DNA PLANT TECHNOLOGY CORP.
XX
PI   Oeller P,  Gutterson N;
XX
DR   WPI; 2001-025165/03.
DR   P-PSDB; AAB50265.
XX
PT   Novel nucleic acid involved in controlling plant flowering processes is
PT   useful for generating transgenic plants, in particular strawberry
PT   plants having altered flowering behavior such as early, delayed or
PT   day-neutral flowering  -
XX
PS   Claim 3; Page 57-58; 97pp; English.
XX
CC   The present invention provides the nucleic acid and protein sequences of
CC   a number of proteins from the strawberry which are involved in the
CC   regulation of flowering. These were identified using primers based on the
CC   homologous sequences from A. thaliana, B. napus and R. sativus. They can
CC   be used in the production of transgenic field crops whose flowering is
CC   regulated and the time of fruiting and harvesting can be manipulated.
XX
SQ   Sequence 738 BP; 227 A; 141 C; 197 G; 173 T; 0 other;

  Query Match             28.8%;  Score 305.8;  DB 22;  Length 738;
  Best Local Similarity   73.4%;  Pred. No. 1.6e-68;
  Matches  391;  Conservative    0;  Mismatches  142;  Indels    0;  Gaps    0;
```

```
Qy    101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
          ||||||| ||||||||||||||||||||||||||||||||||||||| ||||| || ||
Db      1 ATGGGAAGGGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAACAGGCAGGTGACG 60

Qy    161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
          || || ||  ||||  |||| ||||| ||||| || ||||||||||||||  | ||||||
Db     61 TTTTCGAAACGAAGAACTGGGTTGCTGAAGAAGGCGCATGAGATCTCTGTCATGTGCGAT 120

Qy    221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
          ||| |||||||| | || |||||||||  |||||| || ||||| |||||  | || ||
Db    121 GCTCAGGTTGCTTTGATTGTCTTCTCTAACAAAGGGAAGCTCTTTGAATACGCAACTGAT 180

Qy    281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
          || |||||||||   ||||||||||||||||  | || |  ||| | || | ||| ||
Db    181 TCATGCATGGAGGATATACTTGAACGCTATGAGAGGTACTCGTATGCGGAGAGACAGCTA 240

Qy    341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
          || |  |   |  ||     |   |  || |||    | ||||||||||| | | ||||
Db    241 GTGGAACCTAATTTTGATTCACAGGGTAACTGGCCCTTCGAACATGCTAGGTTTAAGGTG 300

Qy    401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
          | ||||||| | ||  |||  |||  |||| | | | |||| |||||||| ||||| |||
Db    301 AAAGTTGAGCTTCTGCAGAGAAACCTAAGGCACTATTTGGGAGAAGATCTGGATTCATTG 360

Qy    461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
          ||  | || ||| |||||||| |||||||| ||||| ||  | ||| | |||   ||| |
Db    361 AGTATCAAAGAGATCCAAAGTTTGGAGCAACAGCTTGAAACTGCTCTTAAGCAAATTCGA 420

Qy    521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
          |||||||| ||||||   |||   || || || |||| ||| |||| |||||| || ||
Db    421 TCAAGAAAAAACCAACTCATGCATGAGTCGATCTCTGAGCTACAGAGGAAGGAGAAGGCA 480

Qy    581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAA 633
          |  | || || ||||| | ||||   || ||||||||||||   ||||||||
Db    481 ATCAAGGAGCAGAACAACTTGCTTTCAAAGAAGATTAAGGAGCACGAGAAGAA 533
```

RESULT 11
AAV58316
ID AAV58316 standard; DNA; 794 BP.
XX
AC AAV58316;
XX
DT 19-NOV-1998 (first entry)
XX
DE Brassica oleracea AP1 gene.
XX
KW CAULIFLOWER gene; CAL gene; shoot meristem conversion; floral meristem;
KW early flowering promotion; angiosperm; AP1 gene; ds.
XX
OS Brassica oleracea.
XX
FH Key Location/Qualifiers
FT CDS 36..794
FT /*tag= a
XX
PN US5811536-A.
XX
PD 22-SEP-1998.
XX
PF 26-JAN-1996; 96US-0592214.
XX
PR 26-JAN-1996; 96US-0592214.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Yanofsky MF;
XX
DR WPI; 1998-530945/45.
DR P-PSDB; AAW69329.
XX
PT Cloned CAULIFLOWER genes - and vectors for converting shoot
PT meristems to floral meristems
XX
PS Disclosure; Fig 2; 93pp; English.
XX
CC This sequence encodes the Brassica oleracea AP1 protein. This sequence
CC was used to isolate DNA encoding the Arabidopsis thaliana CAULIFLOWER
CC (CAL) protein of the invention. An expression vector containing the CAL
CC DNA sequence can be used to convert shoot meristems to floral meristems,
CC especially to promote early flowering in angiosperms.
XX
SQ Sequence 794 BP; 255 A; 171 C; 187 G; 181 T; 0 other;

Query Match 28.3%; Score 301; DB 19; Length 794;
Best Local Similarity 68.2%; Pred. No. 2.8e-67;
Matches 418; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```
Qy     85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
          || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db     20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy    145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
          |||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||||
Db     80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy    205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
          |||||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db    140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy    265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           ||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  ||
Db    200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy    325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
            | || | ||| ||| | |  |  ||          | |   || |||    | ||  |
Db    260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy    385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
          |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db    320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy    445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
          |||  |  |  | |||||   ||||| ||||| |   | ||||| ||||| ||| | ||
Db    380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy    505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
          | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db    440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy    565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  ||||| |||||| | || || || |||    ||||  |||  ||||||||||||
Db    500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy    625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
          ||| ||        |   |||| |||  ||||  | | |      | |    |   |
Db    560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy    685 TCTTCTGCCTCAA 697
            | || ||||| |
Db    620 GCCTCCGCCTCCA 632
```

RESULT 12
AAT86629
ID AAT86629 standard; DNA; 794 BP.
XX
AC AAT86629;
XX
DT 18-JUN-1998 (first entry)
XX
DE APETALA1 gene from Brassica oleracea.
XX
KW Transgenic plant; ectopically expressed; meristem gene; APETALA1;
KW AP1; CAULIFLOWER; CAL; LEAFY; LFY; reproductive development;

```
KW   angiosperm; ds.
XX
OS   Brassica oleracea.
XX
FH   Key             Location/Qualifiers
FT   CDS             36..794
FT                   /*tag=  a
FT                   /product= AP1_gene_product
FT                   /note= "No stop codon shown"
XX
PN   WO9746077-A1.
XX
PD   11-DEC-1997.
XX
PF   05-JUN-1996;   96WO-US09429.
XX
PR   05-JUN-1996;   96WO-US09429.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Weigel D,  Yanofsky MF;
XX
DR   WPI; 1998-041769/04.
DR   P-PSDB; AAW43110.
XX
PT   Transgenic plant comprising an ectopically expressed floral meristem
PT   gene  - for promoting early reproductive development and controlling
PT   the time of seed-derived crop harvest(s) in e.g. grapes, beans,
PT   corn, wheat, etc
XX
PS   Disclosure; Pages 121-122; 158pp; English.
XX
CC   This sequence encodes a floral meristem identity gene product,
CC   APETALA1 (AP1) from Brassica oleracea. The invention relates to
CC   a non-naturally occurring seed plant comprising a first ectopically
CC   expressible nucleic acid encoding a first floral meristem identity gene
CC   product, provided that the first nucleic acid is not ectopically
CC   expressed due to a mutation in an endogenous TERMINAL FLOWER gene.
CC   The invention describes a method of converting shoot meristem to floral
CC   meristem, especially in order to promote early reproductive development,
CC   in an angiosperm, which comprises introducing a first ectopically
CC   expressible nucleic acid molecule encoding a first floral meristem
CC   identity gene product into the angiosperm. Ectopic expression of the
CC   floral meristem gene product in the shoot meristem tissue of the
CC   transgenic plant, allows selection of the time of seed development in
CC   the plant which can be useful for manipulating the time of crop harvest
CC   in seed-derived crops such as grapes, beans, corn, wheat, rice, hop,
CC   etc., or to hasten the breeding of e.g. tree species, for insect or
CC   disease resistance.
XX
SQ   Sequence 794 BP; 255 A; 171 C; 187 G; 181 T; 0 other;

  Query Match             28.3%;  Score 301;  DB 19;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 2.8e-67;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           ||||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           |||||  ||||| || ||||| ||||||| ||||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
           | || | ||| ||| | |  |  ||          | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
           |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db     320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy     445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
           |||  |  |  |  |||||   ||||| ||||| |   | |||||  |||| ||| | ||
Db     380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy     505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
           | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db     440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy     565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  ||||| |||||| | || || || ||||    ||||  |||  ||||||||||||
Db     500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           ||| ||        |   |||| |||  ||||  | | |      | |      |   |
Db     560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy     685 TCTTCTGCCTCAA 697
           | || |||||  |
Db     620 GCCTCCGCCTCCA 632


RESULT 13
AAV02761
ID   AAV02761 standard; cDNA; 794 BP.
XX
AC   AAV02761;
XX
DT   08-JUN-1998  (first entry)
XX
DE   Brassica oleracea floral meristem identity gene APETALA1 cDNA.
XX
KW   Floral meristem identity gene; APETALA1; AP1; transgenic plant;
KW   angiosperm; seed development; ds.
XX
OS   Brassica oleracea.
XX
FH   Key             Location/Qualifiers
FT   CDS             36..794
FT                   /*tag=  a
XX
PN   WO9746078-A1.
XX
PD   11-DEC-1997.
XX
PF   05-JUN-1996;   96WO-US09453.
XX
PR   05-JUN-1996;   96WO-US09453.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Yanofsky MF;
XX
DR   WPI; 1998-041770/04.
DR   P-PSDB; AAW39132.
XX
PT   Nucleic acids encoding AP1 floral meristem identity gene product
PT   from cauliflower and maize - useful for promoting early reproductive
PT   development and controlling the time of seed-derived crop harvest in
PT   e.g. grapes, beans, corn, wheat, etc
XX
PS   Claim 2; Page 121-122; 156pp; English.
XX
CC   This sequence comprises Brassica oleracea APETALA (AP1) cDNA,
CC   which encodes a floral meristem identity gene product (see AAW39132)
CC   that is involved in the conversion of shoot meristem to floral
CC   meristem.  Mutation of the AP1 gene results in replacement of basal
```

```
CC   flowers by inflorescence shoots that are not subtended by flowers.
CC   The invention relates to an expression vector comprising a nucleic
CC   acid sequence encoding a floral meristem gene product operably
CC   linked to a heterologous regulatory element (see AAV02770-75).  The
CC   floral meristem gene product is selected from AP1, LFY (LEAFY) and
CC   CAL (CAULIFLOWER) (see AAV02760-67).  The expression vector is used to
CC   convert shoot meristem tissue to floral meristem tissue in
CC   transgenic plants, especially angiosperms or gymnosperms, thereby
CC   promoting early reproductive development in these plants.  This can
CC   be used to manipulate the time of crop harvest and to hasten
CC   breeding time.
XX
SQ   Sequence 794 BP; 255 A; 171 C; 187 G; 181 T; 0 other;

  Query Match             28.3%;  Score 301;  DB 19;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 2.8e-67;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || ||||| ||||| |  | ||||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           ||||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           ||||| ||||| || ||||| ||||||| |||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
             | || | ||| ||| | |  |  ||        | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
           |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db     320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy     445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
           |||  |  |  |  ||||   ||||| ||||| |   | ||||| ||||| ||| | ||
Db     380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy     505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
           | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db     440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy     565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  ||||| |||||| | || || || ||||    ||||  |||  ||||||||||||
Db     500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           ||| ||         |   |||| ||| |||| | | |       | |       |    |
Db     560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy     685 TCTTCTGCCTCAA 697
           | || ||||| |
Db     620 GCCTCCGCCTCCA 632


RESULT 14
AAV06019
ID   AAV06019 standard; cDNA; 794 BP.
XX
AC   AAV06019;
XX
DT   08-JUN-1998  (first entry)
XX
DE   Brassica floral meristem identity gene APETALA1 cDNA.
XX
KW   Floral meristem identity gene; APETALA1; AP1; transgenic plant;
KW   angiosperm; seed development; ds.
XX
OS   Brassica oleracea.
XX
FH   Key             Location/Qualifiers
FT   CDS             36..794
FT                   /*tag=  a
XX
PN   WO9746079-A1.
XX
PD   11-DEC-1997.
XX
PF   04-JUN-1997;   97WO-US09682.
XX
PR   05-JUN-1996;   96US-0659188.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Yanofsky MF;
XX
DR   WPI; 1998-041771/04.
DR   P-PSDB; AAW43327.
XX
PT   Use of floral meristem identity genes to convert shoot meristem to
PT   floral meristem - for promoting early reproductive development in
PT   the plants and control timing of seed-derived crop harvest(s) in
PT   e.g. grapes, beans, corn and wheat
XX
PS   Claim 6; Page 106-107; 147pp; English.
XX
CC   This sequence comprises Brassica oleracea APETALA (AP1) cDNA,
CC   which encodes a floral meristem identity gene product (see AAW43327)
CC   that is involved in the conversion of shoot meristem to floral
CC   meristem.  Mutation of the AP1 gene results in replacement of basal
CC   flowers by inflorescence shoots that are not subtended by flowers.
CC   The invention relates to a recombinant nucleic acid molecule that
CC   comprises an inducible regulatory element (see AAV06025-28) operably
CC   linked to a nucleic acid molecule (see AAV06018-24) encoding a floral
CC   meristem identity gene product (see AAW43326-32), especially AP1,
CC   CAULIFLOWER (CAL) or LEAFY (LFY).  The nucleic acids are used to
CC   convert shoot meristem tissue into floral meristem tissue in
CC   transgenic plants, thereby promoting early reproductive development
CC   in these plants.  This can be useful for manipulating the time of
CC   crop harvest in seed-derived crops and to hasten breeding time.
XX
SQ   Sequence 794 BP; 255 A; 171 C; 187 G; 181 T; 0 other;

  Query Match             28.3%;  Score 301;  DB 19;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 2.8e-67;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || ||||| ||||| |  | ||||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           ||||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           ||||| ||||| || ||||| ||||||| |||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
             | || | ||| ||| | |  |  ||        | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
```

```
Db    320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy    445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
Db    380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy    505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
Db    440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy    565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
Db    500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy    625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
Db    560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy    685 TCTTCTGCCTCAA 697
Db    620 GCCTCCGCCTCCA 632
```

RESULT 15
AAC61408
ID AAC61408 standard; cDNA; 794 BP.
XX
AC AAC61408;
XX
DT 19-FEB-2001 (first entry)
XX
DE cDNA encoding a APETALA1 (AP1) polypeptide.
XX
KW APETALA1; AP1; floral meristem identity gene; CAULIFLOWER; CAL; LEAFY;
KW LFY; floral meristem; early flowering; ds.
XX
OS Brassica oleracea.
XX

Key	Location/Qualifiers
FH	
FT CDS	36..794
FT	/*tag= a
FT	/product= "APETALA1"
FT	/note= "no stop codon given"

XX
PN US6127123-A.
XX
PD 03-OCT-2000.
XX
PF 09-SEP-1998; 98US-0149976.
XX
PR 26-JAN-1996; 96US-0592214.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Yanofsky MF;
XX
DR WPI; 2000-618379/59.
DR P-PSDB; AAB19242.
XX
PT Identifying a Brassica having a cauliflower phenotype involves
PT detecting a polymorphism associated with cauliflower locus comprising a
PT modified cauliflower allele that does not encode active cauliflower
PT gene product -
XX
PS Example 1; Fig 2A-B; 93pp; English.
XX
CC The present sequence encodes an APETALA1 (AP1) polypeptide. The AP1
CC polypeptide is an ectopically expressible floral meristem identity
CC gene product. The specification also describes CAULIFLOWER (CAL) and
CC LEAFY (LFY) gene products. CAL is involved in the conversion of shoot
CC meristem to floral meristem. CAL is highly conserved among different
CC angiosperms. The CAL polynucleotides may be used to shoot meristem to
CC floral meristem, and to promote early flowering in angiosperms.
XX
SQ Sequence 794 BP; 255 A; 171 C; 187 G; 181 T; 0 other;

Query Match 28.3%; Score 301; DB 21; Length 794;
Best Local Similarity 68.2%; Pred. No. 2.8e-67;
Matches 418; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```
Qy     85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
Db     20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy    145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
Db     80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy    205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
Db    140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy    265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
Db    200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy    325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
Db    260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy    385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
Db    320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy    445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
Db    380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy    505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
Db    440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy    565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
Db    500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy    625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
Db    560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy    685 TCTTCTGCCTCAA 697
Db    620 GCCTCCGCCTCCA 632
```

Search completed: February 2, 2003, 05:39:31
Job time : 294 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: February 1, 2003, 23:33:46 ; Search time 59 Seconds
(without alignments)
5520.182 Million cell updates/sec

Title: US-09-981-087A-1
Perfect score: 1062
Sequence: 1 cccagagagacataagaaag.....taatgatttctcatgaaata 1062

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1062	100.0	1062	4	US-09-067-800-1	Sequence 1, Appli
2	1062	100.0	1062	4	US-09-105-652-1	Sequence 1, Appli
3	1062	100.0	1062	4	US-09-349-677-1	Sequence 1, Appli
4	301	28.3	794	1	US-08-592-214A-3	Sequence 3, Appli
5	301	28.3	794	3	US-08-659-188-3	Sequence 3, Appli
6	301	28.3	794	3	US-08-655-227-3	Sequence 3, Appli
7	301	28.3	794	3	US-08-655-241-3	Sequence 3, Appli
8	301	28.3	794	3	US-09-149-976-3	Sequence 3, Appli
9	301	28.3	794	4	US-09-398-326-3	Sequence 3, Appli
10	296.8	27.9	1054	2	US-08-576-156-1	Sequence 1, Appli
11	296.8	27.9	1057	3	US-08-659-188-1	Sequence 1, Appli
12	296.8	27.9	1057	3	US-08-655-227-1	Sequence 1, Appli
13	296.8	27.9	1057	3	US-08-655-241-1	Sequence 1, Appli
14	296.8	27.9	1057	4	US-09-398-326-1	Sequence 1, Appli
15	296.8	27.9	1215	1	US-08-592-214A-1	Sequence 1, Appli
16	296.8	27.9	1215	3	US-09-149-976-1	Sequence 1, Appli
17	291.4	27.4	768	1	US-08-592-214A-5	Sequence 5, Appli
18	291.4	27.4	768	3	US-08-659-188-5	Sequence 5, Appli
19	291.4	27.4	768	3	US-08-655-227-5	Sequence 5, Appli
20	291.4	27.4	768	3	US-08-655-241-5	Sequence 5, Appli
21	291.4	27.4	768	3	US-09-149-976-5	Sequence 5, Appli
22	291.4	27.4	768	4	US-09-398-326-5	Sequence 5, Appli
23	266	25.0	779	1	US-08-592-214A-9	Sequence 9, Appli
24	266	25.0	779	3	US-08-659-188-9	Sequence 9, Appli
25	266	25.0	779	3	US-08-655-227-9	Sequence 9, Appli
26	266	25.0	779	3	US-08-655-241-9	Sequence 9, Appli
27	266	25.0	779	3	US-09-149-976-9	Sequence 9, Appli
28	266	25.0	779	4	US-09-398-326-9	Sequence 9, Appli
29	255.8	24.1	1345	1	US-08-592-214A-7	Sequence 7, Appli
30	255.8	24.1	1345	3	US-08-659-188-7	Sequence 7, Appli
31	255.8	24.1	1345	3	US-08-655-227-7	Sequence 7, Appli
32	255.8	24.1	1345	3	US-08-655-241-7	Sequence 7, Appli
33	255.8	24.1	1345	3	US-09-149-976-7	Sequence 7, Appli
34	255.8	24.1	1345	4	US-09-398-326-7	Sequence 7, Appli
35	252.6	23.8	756	1	US-08-592-214A-13	Sequence 13, Appl
36	252.6	23.8	756	3	US-08-659-188-13	Sequence 13, Appl
37	252.6	23.8	756	3	US-08-655-227-13	Sequence 13, Appl
38	252.6	23.8	756	3	US-08-655-241-13	Sequence 13, Appl
39	252.6	23.8	756	3	US-09-149-976-13	Sequence 13, Appl
40	252.6	23.8	756	4	US-09-398-326-13	Sequence 13, Appl
41	247.8	23.3	756	1	US-08-592-214A-11	Sequence 11, Appl
42	247.8	23.3	756	3	US-08-659-188-11	Sequence 11, Appl
43	247.8	23.3	756	3	US-08-655-227-11	Sequence 11, Appl
44	247.8	23.3	756	3	US-08-655-241-11	Sequence 11, Appl
45	247.8	23.3	756	3	US-09-149-976-11	Sequence 11, Appl

ALIGNMENTS

```
RESULT 1
US-09-067-800-1
; Sequence 1, Application US/09067800
; Patent No. 6198024
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Ferrandiz, Cristina
;    TITLE OF INVENTION:  Seed Plants Characterized by Delayed
;    TITLE OF INVENTION:  Seed Dispersal
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/067,800
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 2948
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1062 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  101..827
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1062
;      OTHER INFORMATION:  /note= "There is a poly(A) tail at
;      OTHER INFORMATION:  the end."
;    FEATURE:
;      NAME/KEY:  misc_feature
```

```
;       LOCATION:  1..1062
;       OTHER INFORMATION:  /note= "Nucleotide and Deduced
;       OTHER INFORMATION:  Amino Acid Sequences of the AGL8 cDNA clone."
US-09-067-800-1

  Query Match             100.0%;  Score 1062;  DB 4;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-299;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
           |||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 2
US-09-105-652-1
; Sequence 1, Application US/09105652
; Patent No. 6229068
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT: Martienssen, Robert
;  APPLICANT: Ferrandiz, Cristina
;  APPLICANT: Gu, Qing
;  TITLE OF INVENTION: Method of Increasing Fruit Size in a Plant
;  FILE REFERENCE: P-UD 3040
;  CURRENT APPLICATION NUMBER: US/09/105,652
;  CURRENT FILING DATE: 1998-06-26
;  EARLIER APPLICATION NUMBER: 60/051,030
;  EARLIER FILING DATE: 1997-06-27
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Arabidopsis sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (101)..(826)
;   FEATURE:
;   NAME/KEY: polyA_signal
;   LOCATION: (1061)
US-09-105-652-1

  Query Match             100.0%;  Score 1062;  DB 4;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-299;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
```

```
Qy    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
          |||||||||||||||||||||||||||||||||||||||||
Db   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
```

```
RESULT 3
US-09-349-677-1
; Sequence 1, Application US/09349677
; Patent No. 6288305
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Ferrandiz, Cristina
;    TITLE OF INVENTION:  Seed Plants Characterized by Delayed
;    TITLE OF INVENTION:  Seed Dispersal
;    NUMBER OF SEQUENCES:  24
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/349,677
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  09/067,800
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 2948
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1062 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  101..827
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1062
;      OTHER INFORMATION:  /note= "There is a poly(A) tail at
;      OTHER INFORMATION:  the end."
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1062
;      OTHER INFORMATION:  /note= "Nucleotide and Deduced
;      OTHER INFORMATION:  Amino Acid Sequences of the AGL8 cDNA clone."
US-09-349-677-1

  Query Match             100.0%;  Score 1062;  DB 4;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-299;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy     61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy    121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy    181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy    241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
           |||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 4
US-08-592-214A-3
; Sequence 3, Application US/08592214A
; Patent No. 5811536
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identify
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,214A
;      FILING DATE:  26-JAN-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1927
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  794 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36..794
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..794
;      OTHER INFORMATION:  /note= "product = Brassica oleracea
;      OTHER INFORMATION:  AP1"
US-08-592-214A-3

  Query Match             28.3%;  Score 301;  DB 1;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 4.2e-78;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           ||||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           |||||  ||||| || ||||| |||||| ||||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
           | || | ||| ||| | |  |  ||          | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
           |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db     320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy     445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
           |||  |  |  |  |||||   ||||| |||||  |    | |||||  |||||  ||| | ||
Db     380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy     505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
           | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db     440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy     565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  ||||| |||||| | || || || ||||    ||||  |||  ||||||||||||||
Db     500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           ||| ||         |   |||| |||  ||||  | | |       | |       |   |
Db     560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy     685 TCTTCTGCCTCAA 697
           | || ||||| |
Db     620 GCCTCCGCCTCCA 632


RESULT 5
US-08-659-188-3
; Sequence 3, Application US/08659188
; Patent No. 6002069
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Seed Plants Exhibiting Inducible Early
;    TITLE OF INVENTION:  Reproductive Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
```

```
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/659,188
;       FILING DATE:  05-JUN-1996
;       CLASSIFICATION:  800
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Campbell, Cathryn A.
;       REGISTRATION NUMBER:  31,815
;       REFERENCE/DOCKET NUMBER:  P-UD 1946
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 535-9001
;       TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  794 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  36..794
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  1..794
;       OTHER INFORMATION:  /note= "product = Brassica oleracea
;       OTHER INFORMATION:  AP1."
US-08-659-188-3

  Query Match             28.3%;  Score 301;  DB 3;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 4.2e-78;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           |||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           ||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
           | || | ||| ||| | |  |  ||          | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
           |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db     320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy     445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
           |||  |  |  | |||||  ||||| |||||  |   | ||||| ||||| ||| | ||
Db     380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy     505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
           | | |||  |||  | || ||||| ||||||  |||||| ||| |||||   || ||||||
Db     440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy     565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  |||||  |||||| | || || || ||||    ||||  |||  ||||||||||||
Db     500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           ||| ||        |  |||| |||  ||||  | | |      | |      |   |
Db     560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy     685 TCTTCTGCCTCAA 697
           | || ||||| |
Db     620 GCCTCCGCCTCCA 632


RESULT 6
US-08-655-227-3
; Sequence 3, Application US/08655227
; Patent No. 6025483
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Maize and Cauliflower APETALA1 Gene
;    TITLE OF INVENTION:  Products and Nucleic Acid Molecules Encoding Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,227
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 2143
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  794 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36..794
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..794
;      OTHER INFORMATION:  /note= "product = Brassica oleracea
;      OTHER INFORMATION:  AP1."
US-08-655-227-3

  Query Match             28.3%;  Score 301;  DB 3;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 4.2e-78;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           |||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199
```

```
Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           ||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
           | || | ||| ||| | |  |  ||          | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
           |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db     320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy     445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
           |||  |  |  |  |||||   ||||| ||||| |   | ||||| ||||| ||| | ||
Db     380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy     505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
           | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db     440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy     565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  ||||| |||||| | || || || ||||    ||||  |||  ||||||||||||||
Db     500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           |||  ||        |   |||| |||  ||||  | | |       | |     |    |
Db     560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy     685 TCTTCTGCCTCAA 697
           |  || |||||  |
Db     620 GCCTCCGCCTCCA 632


RESULT 7
US-08-655-241-3
; Sequence 3, Application US/08655241
; Patent No. 6025543
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Weigel, Detlef
;    TITLE OF INVENTION:  Seed Plants Exhibiting Early Reproductive
;    TITLE OF INVENTION:  Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,241
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  CLASS 800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1894
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  794 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36..794
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..794
;      OTHER INFORMATION:  /note= "product = Brassica oleracea
;      OTHER INFORMATION:  AP1."
US-08-655-241-3

  Query Match             28.3%;  Score 301;  DB 3;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 4.2e-78;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || ||||| |||||  | | ||||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           ||||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           ||||| ||||| || ||||| |||||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
           | || | ||| ||| | |  |  ||          | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
           |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db     320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy     445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
           |||  |  |  |  |||||   ||||| ||||| |   | ||||| ||||| ||| | ||
Db     380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy     505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
           | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db     440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy     565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  ||||| |||||| | || || || ||||    ||||  |||  ||||||||||||||
Db     500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           |||  ||        |   |||| |||  ||||  | | |       | |     |    |
Db     560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy     685 TCTTCTGCCTCAA 697
           |  || |||||  |
Db     620 GCCTCCGCCTCCA 632


RESULT 8
US-09-149-976-3
; Sequence 3, Application US/09149976
; Patent No. 6127123
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identity
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell & Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
```

```
;       COUNTRY:  United States
;       ZIP:  92122
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/149,976
;       FILING DATE:  09-SEP-1998
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/592,214
;       FILING DATE:  26-JAN-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Campbell, Cathryn A.
;       REGISTRATION NUMBER:  31,815
;       REFERENCE/DOCKET NUMBER:  P-UD 3291
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 535-9001
;       TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  794 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  36..794
;     FEATURE:
;       NAME/KEY:  misc_feature
;       LOCATION:  1..794
;       OTHER INFORMATION:  /note= "product = Brassica oleracea
;       OTHER INFORMATION:  AP1"
US-09-149-976-3

  Query Match             28.3%;  Score 301;  DB 3;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 4.2e-78;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || ||||| ||||| |  | ||||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           ||||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
           ||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
             | || | ||| ||| | |  |  ||          | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
           |  || ||| ||||| |   |||||  |  | ||||  ||| | ||| | | | | |||||
Db     320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy     445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
           |||  |  |  |  |||||   ||||| ||||| |   | ||||| ||||| ||| | ||
Db     380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy     505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
           | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db     440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy     565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  ||||| |||||| | || || || ||||    ||||  |||  ||||||||||||
Db     500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           ||| ||        |   |||| |||  ||||  | | |      | |      |   |
Db     560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy     685 TCTTCTGCCTCAA 697
           | || ||||| |
Db     620 GCCTCCGCCTCCA 632


RESULT 9
US-09-398-326-3
; Sequence 3, Application US/09398326
; Patent No. 6355863
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Seed Plants Exhibiting Inducible Early
;    TITLE OF INVENTION:  Reproductive Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/398,326
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/659,188
;      FILING DATE:  05-JUN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 3739
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  794 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36..794
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..794
;      OTHER INFORMATION:  /note= "product = Brassica oleracea
;      OTHER INFORMATION:  AP1."
US-09-398-326-3

  Query Match             28.3%;  Score 301;  DB 4;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 4.2e-78;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy      85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
           || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db      20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79
```

```
Qy     145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
           |||||| ||||| || ||||| || ||||| ||||| |  | ||||||||||||||||||
Db      80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy     205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
           ||||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db     140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy     265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
            ||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  ||
Db     200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy     325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
             | || | ||| ||| | |  |  ||          | |   || |||    | ||  |
Db     260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy     385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
           |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db     320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy     445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
           |||  |  |  | |||||  ||||| ||||| |   | ||||| ||||| ||| | ||
Db     380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy     505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
           | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db     440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy     565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
           |  ||||| |||||| | || || || ||||    ||||  |||  ||||||||||||
Db     500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           ||| ||        |   |||| |||  ||||  | | |      | |      |    |
Db     560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy     685 TCTTCTGCCTCAA 697
           | || ||||| |
Db     620 GCCTCCGCCTCCA 632


RESULT 10
US-08-576-156-1
; Sequence 1, Application US/08576156
; Patent No. 5844119
;  GENERAL INFORMATION:
;    APPLICANT:  Weigel, Detlef
;    TITLE OF INVENTION:  Genetically Modified Plants Having Modulated
;    TITLE OF INVENTION:  Flower Development
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson P.C.
;      STREET:  4225 Executive Square, Suite 1400
;      CITY:  La Jolla
;      STATE:  CA
;      COUNTRY:  US
;      ZIP:  92037
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows95
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/576,156
;      FILING DATE:  12-21-95
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/360,336
;      FILING DATE:  12-21-94
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Haile, Lisa A.
;      REGISTRATION NUMBER:  38,347
;      REFERENCE/DOCKET NUMBER:  07251/013001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-678-5070
;      TELEFAX:  619-678-5099
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1054
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-576-156-1

  Query Match             27.9%;  Score 296.8;  DB 2;  Length 1054;
  Best Local Similarity   69.5%;  Pred. No. 8e-77;
  Matches  403;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy      84 TTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGA 143
           ||  || ||    | | |||||||| ||||||||||  ||||||||||||||||||||
Db     107 TTAAAAAAGGATCAAAAATGGGAAGGGGTAGGGTTCAATTGAAGAGGATAGAGAACAAGA 166

Qy     144 TCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGA 203
           ||||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db     167 TCAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCTCATGAGA 226

Qy     204 TCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCT 263
           ||||||||||||| |||||||| ||||||||  | ||||||||    || || |||||||
Db     227 TCTCTGTTCTCTGTGATGCTGAAGTTGCTCTTGTTGTCTTCTCCCATAAGGGGAAACTCT 286

Qy     264 TCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTAT 323
           ||||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  |
Db     287 TCGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGGTACTCTT 346

Qy     324 ATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAAC 383
           |  | || | ||| ||| |||  |  ||          | |  ||| |||    | ||
Db     347 ACGCCGAAAGACAGCTTATTGCACCTGAGTCCGACGTCAATACAAACTGGTCGATGGAGT 406

Qy     384 ATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGG 443
           ||   | ||| ||||| |   ||||| |  | ||||  ||| | ||| ||| | | ||||
Db     407 ATAACAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCATTATCTTGGGG 466

Qy     444 AAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAG 503
           ||||  |  |  |  |||||   || ||||| || |   ||||||| ||||| ||  | |
Db     467 AAGACTTGCAAGCAATGAGCCCTAAAGAGCTTCAGAATCTGGAGCAGCAGCTTGACACTG 526

Qy     504 CTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCC 563
           || | |||  |||  |  | ||||| ||||||  ||||| ||| |||||   || |||||
Db     527 CTCTTAAGCACATCCGCACTAGAAAAAACCAACTTATGTACGAGTCCATCAATGAGCTCC 586

Qy     564 AGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGA 623
           | || ||||| || ||| | || || || ||||    ||||  ||| |||| |||||||
Db     587 AAAAAAAGGAGAAGGCCATACAGGAGCAAAACAGCATGCTTTCTAAACAGATCAAGGAGA 646

Qy     624 GGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCA 663
           |||| || |      |    ||| | |  || |  |  ||
Db     647 GGGAAAAAATTCTTAGGGCTCAACAGGAGCAGTGGGATCA 686


RESULT 11
US-08-659-188-1
; Sequence 1, Application US/08659188
; Patent No. 6002069
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Seed Plants Exhibiting Inducible Early
;    TITLE OF INVENTION:  Reproductive Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/659,188
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1946
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1057 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  124..893
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1057
;      OTHER INFORMATION:  /note= "product = Arabidopsis
;      OTHER INFORMATION:  thaliana AP1."
US-08-659-188-1

  Query Match             27.9%;  Score 296.8;  DB 3;  Length 1057;
  Best Local Similarity   69.5%;  Pred. No. 8e-77;
  Matches  403;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy      84 TTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGA 143
           ||  || ||    | | |||||||| ||||||||||  ||||||||||||||||||||
Db     107 TTAAAAAAGGATCAAAAATGGGAAGGGGTAGGGTTCAATTGAAGAGGATAGAGAACAAGA 166

Qy     144 TCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGA 203
           ||||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db     167 TCAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCTCATGAGA 226

Qy     204 TCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCT 263
           |||||||||||||| |||||||| ||||||||  | ||||||||    || || |||||||
Db     227 TCTCTGTTCTCTGTGATGCTGAAGTTGCTCTTGTTGTCTTCTCCCATAAGGGGAAACTCT 286

Qy     264 TCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTAT 323
           ||||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  |
Db     287 TCGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGGTACTCTT 346

Qy     324 ATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAAC 383
           |  | || | ||| ||| ||| |  ||          | |  ||| |||    | ||
Db     347 ACGCCGAAAGACAGCTTATTGCACCTGAGTCCGACGTCAATACAAACTGGTCGATGGAGT 406

Qy     384 ATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGG 443
           ||   | ||| ||||| |   ||||| |  | ||||  ||| | ||| ||| | | ||||
Db     407 ATAACAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCATTATCTTGGGG 466

Qy     444 AAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAG 503
           ||||  |  |  |  ||||    || ||||| || |   ||||||| ||||| ||  | |
Db     467 AAGACTTGCAAGCAATGAGCCCTAAAGAGCTTCAGAATCTGGAGCAGCAGCTTGACACTG 526

Qy     504 CTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCC 563
           || | |||  |||  |  | ||||| ||||||  ||||| ||| |||||   || ||||
Db     527 CTCTTAAGCACATCCGCACTAGAAAAAACCAACTTATGTACGAGTCCATCAATGAGCTCC 586

Qy     564 AGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGA 623
           | || ||||| || ||| | || || || ||||    ||||   ||| |||| |||||||
Db     587 AAAAAAAGGAGAAGGCCATACAGGAGCAAAACAGCATGCTTTCTAAACAGATCAAGGAGA 646

Qy     624 GGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCA 663
```

```
           |||| || |      |    ||| | |  || |  |  ||
Db     647 GGGAAAAAATTCTTAGGGCTCAACAGGAGCAGTGGGATCA 686


RESULT 12
US-08-655-227-1
; Sequence 1, Application US/08655227
; Patent No. 6025483
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Maize and Cauliflower APETALA1 Gene
;    TITLE OF INVENTION:  Products and Nucleic Acid Molecules Encoding Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,227
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 2143
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1057 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  124..893
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1057
;      OTHER INFORMATION:  /note= "product = Arabidopsis
;      OTHER INFORMATION:  thaliana AP1."
US-08-655-227-1

  Query Match             27.9%;  Score 296.8;  DB 3;  Length 1057;
  Best Local Similarity   69.5%;  Pred. No. 8e-77;
  Matches  403;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy      84 TTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGA 143
           ||  || ||    | | |||||||| ||||||||||  ||||||||||||||||||||
Db     107 TTAAAAAAGGATCAAAAATGGGAAGGGGTAGGGTTCAATTGAAGAGGATAGAGAACAAGA 166

Qy     144 TCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGA 203
           ||||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db     167 TCAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCTCATGAGA 226

Qy     204 TCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCT 263
           |||||||||||||| |||||||| ||||||||  | ||||||||    || || |||||||
Db     227 TCTCTGTTCTCTGTGATGCTGAAGTTGCTCTTGTTGTCTTCTCCCATAAGGGGAAACTCT 286

Qy     264 TCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTAT 323
           ||||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  |
Db     287 TCGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGGTACTCTT 346

Qy     324 ATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAAC 383
```

```
Db    347 ACGCCGAAAGACAGCTTATTGCACCTGAGTCCGACGTCAATACAAACTGGTCGATGGAGT 406

Qy    384 ATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGG 443
          ||   | ||| ||||| |   ||||| |  | |||| ||| | ||| ||| | | ||||
Db    407 ATAACAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCATTATCTTGGGG 466

Qy    444 AAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAG 503
          ||||  |  |  |  |||||   || ||||| || |   ||||||| ||||| ||  | |
Db    467 AAGACTTGCAAGCAATGAGCCCTAAAGAGCTTCAGAATCTGGAGCAGCAGCTTGACACTG 526

Qy    504 CTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCC 563
          || | |||  |||  |  | ||||| ||||||  ||||| ||| |||||   || ||||
Db    527 CTCTTAAGCACATCCGCACTAGAAAAAACCAACTTATGTACGAGTCCATCAATGAGCTCC 586

Qy    564 AGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGA 623
          | || ||||| || ||| | || || || ||||    ||||   ||| |||| |||||||
Db    587 AAAAAAAGGAGAAGGCCATACAGGAGCAAAACAGCATGCTTTCTAAACAGATCAAGGAGA 646

Qy    624 GGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCA 663
          |||| || |      |    ||| | |  || |  |  ||
Db    647 GGGAAAAAATTCTTAGGGCTCAACAGGAGCAGTGGGATCA 686


RESULT 13
US-08-655-241-1
; Sequence 1, Application US/08655241
; Patent No. 6025543
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    APPLICANT:  Weigel, Detlef
;    TITLE OF INVENTION:  Seed Plants Exhibiting Early Reproductive
;    TITLE OF INVENTION:  Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/655,241
;      FILING DATE:  05-JUN-1996
;      CLASSIFICATION:  CLASS 800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1894
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1057 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  124..893
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1057
;      OTHER INFORMATION:  /note= "product = Arabidopsis
;      OTHER INFORMATION:  thaliana AP1."
US-08-655-241-1

  Query Match           27.9%;  Score 296.8;  DB 3;  Length 1057;
  Best Local Similarity  69.5%;  Pred. No. 8e-77;
  Matches  403;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy     84 TTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGA 143
          ||  || ||    | | |||||||| ||||||||||  ||||||||||||||||||||
Db    107 TTAAAAAAGGATCAAAAATGGGAAGGGGTAGGGTTCAATTGAAGAGGATAGAGAACAAGA 166

Qy    144 TCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGA 203
          ||||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db    167 TCAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCTCATGAGA 226

Qy    204 TCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCT 263
          |||||||||||| |||||||| |||||||| | ||||||||    || || |||||||
Db    227 TCTCTGTTCTCTGTGATGCTGAAGTTGCTCTTGTTGTCTTCTCCCATAAGGGGAAACTCT 286

Qy    264 TCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTAT 323
          ||||||| ||||| || ||||| ||||||| |||||||||||||||||  | || |  |
Db    287 TCGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGGTACTCTT 346

Qy    324 ATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAAC 383
          |  | || | ||| ||| |||  |  ||          | |  ||| |||    | ||
Db    347 ACGCCGAAAGACAGCTTATTGCACCTGAGTCCGACGTCAATACAAACTGGTCGATGGAGT 406

Qy    384 ATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGG 443
          ||   | ||| ||||| |   ||||| |  | |||| ||| | ||| ||| | | ||||
Db    407 ATAACAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCATTATCTTGGGG 466

Qy    444 AAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAG 503
          ||||  |  |  |  |||||   || ||||| || |   ||||||| ||||| ||  | |
Db    467 AAGACTTGCAAGCAATGAGCCCTAAAGAGCTTCAGAATCTGGAGCAGCAGCTTGACACTG 526

Qy    504 CTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCC 563
          || | |||  |||  |  | ||||| ||||||  ||||| ||| |||||   || ||||
Db    527 CTCTTAAGCACATCCGCACTAGAAAAAACCAACTTATGTACGAGTCCATCAATGAGCTCC 586

Qy    564 AGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGA 623
          | || ||||| || ||| | || || || ||||    ||||   ||| |||| |||||||
Db    587 AAAAAAAGGAGAAGGCCATACAGGAGCAAAACAGCATGCTTTCTAAACAGATCAAGGAGA 646

Qy    624 GGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCA 663
          |||| || |      |    ||| | |  || |  |  ||
Db    647 GGGAAAAAATTCTTAGGGCTCAACAGGAGCAGTGGGATCA 686


RESULT 14
US-09-398-326-1
; Sequence 1, Application US/09398326
; Patent No. 6355863
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Seed Plants Exhibiting Inducible Early
;    TITLE OF INVENTION:  Reproductive Development and Methods of Making Same
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores LLP
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/398,326
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/659,188
;      FILING DATE:  05-JUN-1996
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 3739
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1057 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  124..893
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1057
;      OTHER INFORMATION:  /note= "product = Arabidopsis
;      OTHER INFORMATION:  thaliana AP1."
US-09-398-326-1

  Query Match             27.9%;  Score 296.8;  DB 4;  Length 1057;
  Best Local Similarity   69.5%;  Pred. No. 8e-77;
  Matches  403;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy      84 TTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGA 143
           ||  || ||    | | ||||||||| ||||||||||  ||||||||||||||||||||
Db     107 TTAAAAAAGGATCAAAAATGGGAAGGGGTAGGGTTCAATTGAAGAGGATAGAGAACAAGA 166

Qy     144 TCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGA 203
           ||||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db     167 TCAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCTCATGAGA 226

Qy     204 TCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCT 263
           |||||||||||||| ||||||||  |||||||| | ||||||||    || || |||||||
Db     227 TCTCTGTTCTCTGTGATGCTGAAGTTGCTCTTGTTGTCTTCTCCCATAAGGGGAAACTCT 286

Qy     264 TCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTAT 323
           ||||||| ||||| || ||||| ||||||| ||||||||||||||||  | || |  |
Db     287 TCGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGGTACTCTT 346

Qy     324 ATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAAC 383
           |  | || | ||| ||| |||  |  ||          | |  ||| |||    | ||
Db     347 ACGCCGAAAGACAGCTTATTGCACCTGAGTCCGACGTCAATACAAACTGGTCGATGGAGT 406

Qy     384 ATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGG 443
           ||   | ||| ||||| |  ||||| |  | ||||  ||| | ||| ||| | | ||||
Db     407 ATAACAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCATTATCTTGGGG 466

Qy     444 AAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAG 503
           ||||  |  | |  ||||| || ||||| || |   ||||||| ||||| ||  | |
Db     467 AAGACTTGCAAGCAATGAGCCCTAAAGAGCTTCAGAATCTGGAGCAGCAGCTTGACACTG 526

Qy     504 CTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCC 563
           || | |||  |||  |  | ||||| ||||||  ||||| ||| |||||   || ||||
Db     527 CTCTTAAGCACATCCGCACTAGAAAAAACCAACTTATGTACGAGTCCATCAATGAGCTCC 586

Qy     564 AGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGA 623
           | || ||||| || ||| | || || || ||||    ||||   ||| |||| |||||||
Db     587 AAAAAAAGGAGAAGGCCATACAGGAGCAAAACAGCATGCTTTCTAAACAGATCAAGGAGA 646

Qy     624 GGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCA 663
           |||| || |      |    ||| | |  || |  |  ||
Db     647 GGGAAAAAATTCTTAGGGCTCAACAGGAGCAGTGGGATCA 686


RESULT 15
US-08-592-214A-1
; Sequence 1, Application US/08592214A
; Patent No. 5811536
;  GENERAL INFORMATION:
;    APPLICANT:  Yanofsky, Martin F.
;    TITLE OF INVENTION:  Cauliflower Floral Meristem Identify
;    TITLE OF INVENTION:  Genes and Methods of Using Same
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/592,214A
;      FILING DATE:  26-JAN-1996
;      CLASSIFICATION:  536
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-UD 1927
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1215 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(141..905, 909..971, 975..1047)
;    FEATURE:
;      NAME/KEY:  misc_feature
;      LOCATION:  1..1215
;      OTHER INFORMATION:  /note= "product = Arabidopsis
;      OTHER INFORMATION:  thaliana AP1"
US-08-592-214A-1

  Query Match             27.9%;  Score 296.8;  DB 1;  Length 1215;
  Best Local Similarity   69.5%;  Pred. No. 8.6e-77;
  Matches  403;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy      84 TTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGA 143
           ||  || ||    | | ||||||||| ||||||||||  ||||||||||||||||||||
Db     124 TTAAAAAAGGATCAAAAATGGGAAGGGGTAGGGTTCAATTGAAGAGGATAGAGAACAAGA 183

Qy     144 TCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGA 203
           ||||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db     184 TCAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCTCATGAGA 243

Qy     204 TCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCT 263
           |||||||||||||| ||||||||  |||||||| | ||||||||    || || |||||||
Db     244 TCTCTGTTCTCTGTGATGCTGAAGTTGCTCTTGTTGTCTTCTCCCATAAGGGAAAACTCT 303

Qy     264 TCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTAT 323
           ||||||| ||||| || ||||| ||||||| ||||||||||||||||  | || |  |
Db     304 TCGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGGTACTCTT 363

Qy     324 ATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAAC 383
           |  | || | ||| ||| |||  |  ||          | |  ||| |||    | ||
Db     364 ACGCCGAAAGACAGCTTATTGCACCTGAGTCCGACGTCAATACAAACTGGTCGATGGAGT 423

Qy     384 ATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGG 443
           ||   | ||| ||||| |  ||||| |  | ||||  ||| | ||| ||| | | ||||
Db     424 ATAACAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCATTATCTTGGGG 483

Qy     444 AAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAG 503
```

```
Db    484 AAGACTTGCAAGCAATGAGCCCTAAAGAGCTTCAGAATCTGGAGCAGCAGCTTGACACTG 543

Qy    504 CTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCC 563

Db    544 CTCTTAAGCACATCCGCACTAGAAAAAACCAACTTATGTACGAGTCCATCAATGAGCTCC 603

Qy    564 AGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGA 623

Db    604 AAAAAAAGGAGAAGGCCATACAGGAGCAAAACAGCATGCTTTCTAAACAGATCAAGGAGA 663

Qy    624 GGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCA 663

Db    664 GGGAAAAAATTCTTAGGGCTCAACAGGAGCAGTGGGATCA 703
```

Search completed: February 2, 2003, 05:34:26
Job time : 64 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: February 2, 2003, 01:40:51 ; Search time 77 Seconds
(without alignments)
6196.354 Million cell updates/sec

Title: US-09-981-087A-1
Perfect score: 1062
Sequence: 1 cccagagagacataagaaag.....taatgatttctcatgaaata 1062

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 396772 seqs, 224632407 residues

Total number of hits satisfying chosen parameters: 793544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1062	100.0	1062	9	US-09-978-729A-1	Sequence 1, Appli
2	1062	100.0	1062	9	US-09-981-087A-1	Sequence 1, Appli
3	1062	100.0	1062	9	US-09-978-382A-1	Sequence 1, Appli
4	1062	100.0	1062	9	US-09-978-740A-1	Sequence 1, Appli
5	1062	100.0	1062	10	US-09-978-730-1	Sequence 1, Appli
6	301	28.3	794	9	US-09-853-450-3	Sequence 3, Appli
7	296.8	27.9	1057	9	US-09-853-450-1	Sequence 1, Appli
8	291.4	27.4	768	9	US-09-853-450-5	Sequence 5, Appli
9	266	25.0	779	9	US-09-853-450-9	Sequence 9, Appli
10	255.8	24.1	1345	9	US-09-853-450-7	Sequence 7, Appli
11	252.6	23.8	756	9	US-09-853-450-13	Sequence 13, Appl
12	247.8	23.3	756	9	US-09-853-450-11	Sequence 11, Appl
13	230	21.7	1241	9	US-09-938-842A-3474	Sequence 3474, Ap
14	194.4	18.3	747	9	US-09-853-450-27	Sequence 27, Appl
15	185.8	17.5	753	9	US-09-853-450-29	Sequence 29, Appl
16	184.6	17.4	777	9	US-09-938-842A-1873	Sequence 1873, Ap
17	158.2	14.9	756	9	US-09-938-842A-718	Sequence 718, App
18	152.4	14.4	714	9	US-09-853-450-31	Sequence 31, Appl
19	149.2	14.0	896	9	US-09-978-729A-3	Sequence 3, Appli
20	149.2	14.0	896	9	US-09-981-087A-3	Sequence 3, Appli
21	149.2	14.0	896	9	US-09-978-382A-3	Sequence 3, Appli
22	149.2	14.0	896	9	US-09-978-740A-3	Sequence 3, Appli
23	149.2	14.0	896	10	US-09-978-730-3	Sequence 3, Appli
24	148.4	14.0	959	9	US-09-978-729A-5	Sequence 5, Appli
25	148.4	14.0	959	9	US-09-981-087A-5	Sequence 5, Appli
26	148.4	14.0	959	9	US-09-978-382A-5	Sequence 5, Appli
27	148.4	14.0	959	9	US-09-978-740A-5	Sequence 5, Appli
28	148.4	14.0	959	10	US-09-978-730-5	Sequence 5, Appli
29	137	12.9	645	9	US-09-938-842A-1776	Sequence 1776, Ap
30	137	12.9	645	9	US-09-853-450-33	Sequence 33, Appl
31	131.8	12.4	705	9	US-09-938-842A-2404	Sequence 2404, Ap
32	128.6	12.1	663	9	US-09-853-450-37	Sequence 37, Appl
33	122.8	11.6	5171	9	US-09-853-450-42	Sequence 42, Appl
34	116.8	11.0	5131	9	US-09-853-450-43	Sequence 43, Appl
35	114.8	10.8	666	9	US-09-938-842A-2442	Sequence 2442, Ap
36	113.8	10.7	5070	9	US-09-853-450-44	Sequence 44, Appl
37	109.4	10.3	5392	9	US-09-853-450-45	Sequence 45, Appl
38	103.4	9.7	633	9	US-09-938-842A-1202	Sequence 1202, Ap
39	102.4	9.6	633	9	US-09-853-450-35	Sequence 35, Appl
40	101	9.5	5483	9	US-09-853-450-47	Sequence 47, Appl
41	99.4	9.4	367	10	US-09-878-574-959	Sequence 959, App
42	95	8.9	1280	10	US-09-970-624-1	Sequence 1, Appli
43	94.2	8.9	5497	10	US-09-916-780A-7	Sequence 7, Appli
44	93	8.8	262	10	US-09-878-574-9686	Sequence 9686, Ap
45	93	8.8	365	10	US-09-878-574-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-978-729A-1
; Sequence 1, Application US/09978729A
; Patent No. US20020178465A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegren, Sarah
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 19452A-000950US
; CURRENT APPLICATION NUMBER: US/09/978,729A
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 60/090,649
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: US 09/339,998
; PRIOR FILING DATE: 1999-06-25
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Arabidopsis sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (101)..(829)
; OTHER INFORMATION: AGAMOUS-LIKE 8 (AGL8)
US-09-978-729A-1

Query Match 100.0%; Score 1062; DB 9; Length 1062;
Best Local Similarity 100.0%; Pred. No. 1.2e-266;
Matches 1062; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
```

```
Db    121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy    181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy    241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
          |||||||||||||||||||||||||||||||||||||||||
Db   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
```

RESULT 2
US-09-981-087A-1
; Sequence 1, Application US/09981087A
; Patent No. US20020178466A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Liljegren, Sarah
; APPLICANT: Farrandiz, Cristina
; APPLICANT: The Regents of the University of California

```
;  TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 19452A-000940US
;  CURRENT APPLICATION NUMBER: US/09/981,087A
;  CURRENT FILING DATE:  2002-05-21
;  PRIOR APPLICATION NUMBER: US 60/090,649
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 09/339,998
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Arabidopsis sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (101)..(829)
;   OTHER INFORMATION: AGAMOUS-LIKE 8 (AGL8)
US-09-981-087A-1

  Query Match             100.0%;  Score 1062;  DB 9;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-266;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy     61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy    121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy    181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy    241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
```

```
Qy    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
          |||||||||||||||||||||||||||||||||||||||||
Db   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 3
US-09-978-382A-1
; Sequence 1, Application US/09978382A
; Publication No. US20020194647A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Liljegren, Sarah
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 19452A-000930US
;  CURRENT APPLICATION NUMBER: US/09/978,382A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 60/090,649
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 09/339,998
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Arabidopsis sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (101)..(829)
;   OTHER INFORMATION: AGAMOUS-LIKE 8 (AGL8)
US-09-978-382A-1

  Query Match           100.0%;  Score 1062;  DB 9;  Length 1062;
  Best Local Similarity  100.0%;  Pred. No. 1.2e-266;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy     61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy    121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy    181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy    241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
          |||||||||||||||||||||||||||||||||||||||||
Db   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 4
US-09-978-740A-1
; Sequence 1, Application US/09978740A
; Publication No. US20030005481A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Liljegren, Sarah
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 19452A-000960US
;  CURRENT APPLICATION NUMBER: US/09/978,740A
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: US 60/090,649
;  PRIOR FILING DATE: 1998-06-25
```

```
;  PRIOR APPLICATION NUMBER: US 09/339,998
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Arabidopsis sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (101)..(829)
;   OTHER INFORMATION: AGAMOUS-LIKE 8 (AGL8)
US-09-978-740A-1

  Query Match             100.0%;  Score 1062;  DB 9;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-266;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
          |||||||||||||||||||||||||||||||||||||||||||
Db    1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 5
US-09-978-730-1
; Sequence 1, Application US/09978730
; Patent No. US20020129403A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Liljegren, Sarah
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Selective Control of Lignin Biosynthesis in Transgenic
;  TITLE OF INVENTION:  Plants
;  FILE REFERENCE: 19452A-000920US
;  CURRENT APPLICATION NUMBER: US/09/978,730
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: US 60/090,649
;  PRIOR FILING DATE: 1998-06-25
;  PRIOR APPLICATION NUMBER: US 09/339,998
;  PRIOR FILING DATE: 1999-06-25
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1062
;   TYPE: DNA
;   ORGANISM: Arabidopsis sp.
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (101)..(829)
;   OTHER INFORMATION: AGAMOUS-LIKE 8 (AGL8)
US-09-978-730-1

  Query Match             100.0%;  Score 1062;  DB 10;  Length 1062;
  Best Local Similarity   100.0%;  Pred. No. 1.2e-266;
  Matches 1062;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCCAGAGAGACATAAGAAAGAAAGAGAGAGAGATACTTTGGTCATTTCAGGGTTGTCG 60

Qy      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCA 120

Qy     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGG 180

Qy     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 TTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGT 240

Qy     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 CTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACT 300

Qy     301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    301 TGAACGCTATGATCGCTATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACA 360

Qy    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 AAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAA 420

Qy    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAG 480

Qy    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTAT 540

Qy    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTC 600

Qy    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGT 660

Qy    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGA 720

Qy    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 TGGCTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAA 780

Qy    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 CTCTCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCAC 840

Qy    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 TCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTT 900

Qy    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 TTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTAT 960

Qy    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTA 1020

Qy   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062
          |||||||||||||||||||||||||||||||||||||||||
Db   1021 AGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAAATA 1062


RESULT 6
US-09-853-450-3
; Sequence 3, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
;   LENGTH: 794
;   TYPE: DNA
;   ORGANISM: Brassica oleracea
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (36)..(794)
;   OTHER INFORMATION: APETALA1 (AP1)
US-09-853-450-3

  Query Match             28.3%;  Score 301;  DB 9;  Length 794;
  Best Local Similarity   68.2%;  Pred. No. 7.6e-69;
  Matches  418;  Conservative    0;  Mismatches  195;  Indels    0;  Gaps    0;

Qy     85 TTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGAT 144
          || || || || | | |||||||| |||||||||||| |||||||||||| ||||||||
Db     20 TTTAAAAGGGATAAAAATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGAT 79

Qy    145 CAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGAT 204
          |||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||||
Db     80 CAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGAT 139

Qy    205 CTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTT 264
          |||||||| || |||||||| ||||| ||  | ||||||||    || || ||||||||
Db    140 CTCTGTTCTGTGTGATGCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTT 199

Qy    265 CGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATA 324
          ||||| |||||| || ||||| ||||||| ||||||||||||||||||  | || |  ||
Db    200 TGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGATACTCTTA 259

Qy    325 TTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACA 384
          | || | ||| ||| | |  |  ||           | |   || |||    | ||  |
Db    260 CGCCGAGAGACAGCTTATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTA 319

Qy    385 TGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGA 444
          |  || ||| ||||| |   ||||| |  | ||||  ||| | ||| | | | | |||||
Db    320 TAATAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGA 379

Qy    445 AGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGC 504
          |||  |  |  |  |||||   ||||| ||||| |   | |||||  ||||| ||| | ||
Db    380 AGACTTGCAAGCAATGAGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGC 439

Qy    505 TATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCA 564
          | | |||  |||  | || ||||| ||||||  ||||| ||| |||||   || ||||||
Db    440 TCTTAAGCACATCCGCTCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCA 499

Qy    565 GAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAG 624
          |  ||||| |||||| | || || || ||||    ||||  |||  ||||||||||||||
Db    500 AAGAAAGGAGAAAGCCATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAG 559

Qy    625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
          ||| ||        |   |||| ||| |||| | | |      | |      |    |
Db    560 GGAAAACGTTCTTAGGGCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATAT 619

Qy    685 TCTTCTGCCTCAA 697
          | || |||||  |
Db    620 GCCTCCGCCTCCA 632


RESULT 7
US-09-853-450-1
; Sequence 1, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 1057
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
```

```
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (124)..(894)
;   OTHER INFORMATION: APETALA1 (AP1)
US-09-853-450-1

  Query Match             27.9%;  Score 296.8;  DB 9;  Length 1057;
  Best Local Similarity   69.5%;  Pred. No. 1.1e-67;
  Matches  403;  Conservative    0;  Mismatches  177;  Indels    0;  Gaps    0;

Qy      84 TTTGAAGAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGA 143
           ||  || ||    | | ||||||| ||||||||||  ||||||||||||||||||||
Db     107 TTAAAAAAGGATCAAAAATGGGAAGGGGTAGGGTTCAATTGAAGAGGATAGAGAACAAGA 166

Qy     144 TCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGA 203
           ||||||| ||||| || ||||| || |||||  ||||| |  | ||||||||||||||||
Db     167 TCAATAGACAAGTGACATTCTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCTCATGAGA 226

Qy     204 TCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCT 263
           |||||||||||| ||||||| |||||||  | ||||||||    || || |||||||
Db     227 TCTCTGTTCTCTGTGATGCTGAAGTTGCTCTTGTTGTCTTCTCCCATAAGGGGAAACTCT 286

Qy     264 TCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTAT 323
           ||||||| ||||| || ||||| ||||||| ||||||||||||||||||  | || |  |
Db     287 TCGAATACTCCACTGATTCTTGTATGGAGAAGATACTTGAACGCTATGAGAGGTACTCTT 346

Qy     324 ATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAAC 383
           |  | || | ||| ||| |||  |  ||          | |  ||| |||    | ||
Db     347 ACGCCGAAAGACAGCTTATTGCACCTGAGTCCGACGTCAATACAAACTGGTCGATGGAGT 406

Qy     384 ATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGG 443
           ||   | ||| ||||| |   ||||| |  | ||||  ||| | ||| ||| | | ||||
Db     407 ATAACAGGCTTAAGGCTAAGATTGAGCTTTTGGAGAGAAACCAGAGGCATTATCTTGGGG 466

Qy     444 AAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAG 503
           ||||  |  |  |  |||||   || ||||| || |   ||||||| ||||| ||  | |
Db     467 AAGACTTGCAAGCAATGAGCCCTAAAGAGCTTCAGAATCTGGAGCAGCAGCTTGACACTG 526

Qy     504 CTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCC 563
           || | |||  |||  |  | ||||| ||||||  ||||| ||| |||||   || |||||
Db     527 CTCTTAAGCACATCCGCACTAGAAAAAACCAACTTATGTACGAGTCCATCAATGAGCTCC 586

Qy     564 AGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGA 623
           | || ||||| || ||| | || || || ||||    ||||   ||| |||| |||||||
Db     587 AAAAAAAGGAGAAGGCCATACAGGAGCAAAACAGCATGCTTTCTAAACAGATCAAGGAGA 646

Qy     624 GGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCA 663
           |||| || |      |    ||| | |  || |  |  ||
Db     647 GGGAAAAAATTCTTAGGGCTCAACAGGAGCAGTGGGATCA 686


RESULT 8
US-09-853-450-5
; Sequence 5, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
;   LENGTH: 768
;   TYPE: DNA
;   ORGANISM: Brassica oleracea var. botrytis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(768)
;   OTHER INFORMATION: APETALA1 (AP1)
US-09-853-450-5

  Query Match             27.4%;  Score 291.4;  DB 9;  Length 768;
  Best Local Similarity   68.0%;  Pred. No. 2.3e-66;
  Matches  406;  Conservative    0;  Mismatches  191;  Indels    0;  Gaps    0;

Qy     101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
           |||||||| |||||||||||| |||||||||||| |||||||||||||| ||||| ||
Db       1 ATGGGAAGGGGTAGGGTTCAGTTGAAGAGGATAGAAAACAAGATCAATAGACAAGTGACA 60

Qy     161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
           ||||| || |||||  ||||| |  | ||||||||||||||||||||||||||| || |||
Db      61 TTCTCGAAAAGAAGAGCTGGTCTTATGAAGAAAGCTCATGAGATCTCTGTTCTGTGTGAT 120

Qy     221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
           ||||| ||||| ||  | ||||||||    || || |||||||| |||||  |||| ||
Db     121 GCTGAAGTTGCGCTTGTTGTCTTCTCCCATAAGGGGAAACTCTTTGAATACCCCACTGAT 180

Qy     281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
           ||||| ||||||  ||||||||||||||||||  | || |  ||  | || | ||| |||
Db     181 TCTTGTATGGAGGAGATACTTGAACGCTATGAGAGATACTCTTACGCCGAGAGACAGCTT 240

Qy     341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
           | | |  ||          | |   || |||    | ||  ||  || ||| |||||
Db     241 ATAGCACCTGAGTCCGACTCCAATACGAACTGGTCGATGGAGTATAATAGGCTTAAGGCT 300

Qy     401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
           |  ||||| |  | ||||  ||| | ||| | | | | |||||||  |  |  |  ||
Db     301 AAGATTGAGCTTTTGGAGAGAAACCAGAGGCACTATCTTGGGGAAGACTTGCAAGCAATG 360

Qy     461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
           |||  ||||| ||||| |  | ||||| |||| | ||| | ||| | ||| |||  |
Db     361 AGCCCTAAGGAACTCCAGAATCTAGAGCAACAGCTTGATACTGCTCTTAAGCACATCCGC 420

Qy     521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
           || ||||| ||||||  ||||| ||| |||||   || |||||| |  ||||| ||||||
Db     421 TCTAGAAAAAACCAACTTATGTACGACTCCATCAATGAGCTCCAAAGAAAGGAGAAAGCC 480

Qy     581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGT 640
           | || || || ||||    ||||  |||  ||||||||||||||| ||        |
Db     481 ATACAGGAACAAAACAGCATGCTTTCCAAGCAGATTAAGGAGAGGGAAAACGTTCTTAGG 540

Qy     641 CAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAA 697
             |||| |||  ||||  | | |       | |       |    | | || ||||| |
Db     541 GCGCAACAAGAGCAATGGGACGAGCAGAACCATGGCCATAATATGCCTCCGCCTCCA 597


RESULT 9
US-09-853-450-9
; Sequence 9, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 779
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (10)..(777)
;   OTHER INFORMATION: CAULIFLOWER (CAL)
```

```
;   NAME/KEY: modified_base
;   LOCATION: (778)..(779)
;   OTHER INFORMATION: n = g, a, c or t
US-09-853-450-9

  Query Match            25.0%;  Score 266;  DB 9;  Length 779;
  Best Local Similarity   69.5%;  Pred. No. 9.5e-60;
  Matches  378;  Conservative    0;  Mismatches  160;  Indels    6;  Gaps    1;

Qy      95 AGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAA 154
           ||||| |||||||| |||||||||  |  ||||||||||||||||||||||||||| |||
Db       4 AGAGAAATGGGAAGGGGTAGGGTTGAATTGAAGAGGATAGAGAACAAGATCAATAGACAA 63

Qy     155 GTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTC 214
           || || ||||| || |||||  ||||| |  | |||||||||| ||||||||||||||
Db      64 GTGACATTCTCGAAAAGAAGAACTGGTCTTTTGAAGAAAGCTCAGGAGATCTCTGTTCTT 123

Qy     215 TGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCC 274
           || ||||| |||||| | || || ||||||||    || |||||| | ||||| || |||
Db     124 TGTGATGCCGAGGTTTCCCTTATTGTCTTCTCCCATAAGGGCAAATTGTTCGAGTACTCC 183

Qy     275 ACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAA 334
           | || ||||||||||||| | |||| |||||||| ||  | ||||  ||  | || | |
Db     184 TCTGAATCTTGCATGGAGAAGGTACTAGAACGCTACGAGAGGTATTCTTACGCCGAGAGA 243

Qy     335 CAACTTGTTGGCCGAGACGTTTCACAAAGTGAA------AATTGGGTTCTAGAACATGCT 388
           || ||  |||  |  |||  |      | || |      || |||    | ||  ||
Db     244 CAGCTGATTGCACCTGACTCTCACGTTAATGCACAGACGAACTGGTCAATGGAGTATAGC 303

Qy     389 AAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGAT 448
           | ||| ||||| |   ||||| |  | ||||  ||| ||||| ||| | |||| |||||
Db     304 AGGCTTAAGGCCAAGATTGAGCTTTTGGAGAGAAACCAAAGGCATTATCTGGGAGAAGAG 363

Qy     449 CTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATC 508
           | ||  |  ||||| | ||||| |||||||   |||||| |||||  ||  | ||| |
Db     364 TTGGAACCAATGAGCCTCAAGGATCTCCAAAATCTGGAGCAGCAGCTTGAGACTGCTCTT 423

Qy     509 AAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAG 568
           |||  |||| | || ||||| || |||   |||   || ||| |       ||||| |
Db     424 AAGCACATTCGCTCCAGAAAAAATCAACTCATGAATGAGTCCCTCAACCACCTCCAAAGA 483

Qy     569 AAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAG 628
           ||||| || |   | || ||  | ||||    ||||  |||| |||| |||||||||||
Db     484 AAGGAGAAGGAGATACAGGAGGAAAACAGCATGCTTACCAAACAGATAAAGGAGAGGGAA 543

Qy     629 AAGA 632
           || |
Db     544 AACA 547


RESULT 10
US-09-853-450-7
; Sequence 7, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 1345
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (149)..(970)
;   OTHER INFORMATION: APETALA1 (AP1)
US-09-853-450-7

  Query Match            24.1%;  Score 255.8;  DB 9;  Length 1345;
  Best Local Similarity   67.6%;  Pred. No. 5.4e-57;
  Matches  359;  Conservative    0;  Mismatches  172;  Indels    0;  Gaps    0;

Qy     101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
           |||||  | || | ||| |||||||| ||||||||||||||| ||  |||| || ||
Db     149 ATGGGGCGCGGCAAGGTACAGCTGAAGCGGATAGAGAACAAGATAAACCGGCAGGTGACC 208

Qy     161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
           ||||| ||| |  ||   ||  |||||||||| || || |||||||| || |||||||||
Db     209 TTCTCCAAGCGCCGGAACGGCCTGCTCAAGAAGGCGCACGAGATCTCCGTCCTCTGCGAT 268

Qy     221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
           || ||||| ||  ||||||||||||  |||| ||||| |||| ||| ||  ||||||||
Db     269 GCCGAGGTCGCCGTCATCGTCTTCTCCCCCAAGGGCAAGCTCTACGAGTACGCCACCGAC 328

Qy     281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
           ||  ||||||| |  || ||||||||||||||| || ||||  ||| | || ||    |||
Db     329 TCCCGCATGGACAAAATTCTTGAACGCTATGAGCGATATTCCTATGCTGAAAAGGCTCTT 388

Qy     341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
           ||       ||   |  |       | |||||||   |  ||| |    || || |||||
Db     389 ATTTCAGCTGAATCTGAAAGTGAGGGAAATTGGTGCCACGAATACAGGAAACTGAAGGCC 448

Qy     401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
           | | |||||    |  | ||   | | | | |  | ||||| || ||||| || || |||
Db     449 AAAATTGAGACCATACAAAAATGCCACAAGCACCTGATGGGAGAGGATCTAGAGTCTTTG 508

Qy     461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
           |     || ||||||||      | ||||| ||||| |||    |  | |||  ||| ||
Db     509 AATCCCAAAGAGCTCCAGCAACTAGAGCAGCAGCTGGATAGCTCACTGAAGCACATCAGA 568

Qy     521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
           ||||| |||| |||   ||||  ||| || || |||| ||| ||||||||||| |   |
Db     569 TCAAGGAAGAGCCACCTTATGGCCGAGTCTATTTCTGAGCTACAGAAGAAGGAGAGGTCA 628

Qy     581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAG 631
           |||| ||  | |||||  | || |  ||  |  ||  ||||||| |||||
Db     629 CTGCAGGAGGAGAACAAGGCTCTGCAGAAGGAACTTGCGGAGAGGCAGAAG 679


RESULT 11
US-09-853-450-13
; Sequence 13, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 756
;   TYPE: DNA
;   ORGANISM: Brassica oleracea var. botrytis
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(453)
;   OTHER INFORMATION: CAULIFLOWER
US-09-853-450-13

  Query Match            23.8%;  Score 252.6;  DB 9;  Length 756;
  Best Local Similarity   66.6%;  Pred. No. 2.8e-56;
  Matches  379;  Conservative    0;  Mismatches  184;  Indels    6;  Gaps    1;
```

```
Qy    101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
          |||||||| |||||||||  |  ||||||||||||||||||||||||  || |||||  ||
Db      1 ATGGGAAGGGGTAGGGTTGAAATGAAGAGGATAGAGAACAAGATCAACAGACAAGTGACG 60

Qy    161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
          || || || |||||  ||||| |  | ||||||| ||||||||||  |||| || |||
Db     61 TTTTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCCCATGAGATCTCGATTCTTTGTGAT 120

Qy    221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
          |||||||||  | || || ||||||||    || || |||||  ||||| || ||  | ||
Db    121 GCTGAGGTTTCCCTTATTGTCTTCTCCCATAAGGGGAAACTGTTCGAGTACTCGTCTGAA 180

Qy    281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
          ||||||||||||| | |||| |||||||| ||  | || |  ||  | || |||||  ||
Db    181 TCTTGCATGGAGAAGGTACTAGAACGCTACGAGAGGTACTCTTACGCCGAGAAACAGCTA 240

Qy    341 GTTGGCCGAGACGTTTCACAAAGTGAA------AATTGGGTTCTAGAACATGCTAAGCTC 394
             |  | ||||  |       | || |       || |||     | ||| ||   | |||
Db    241 AAAGCTCCAGACTCTCACGTCAATGCACAAACGAACTGGTCAATGGAATATAGCAGGCTT 300

Qy    395 AAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGAT 454
          |||||  |   |||||  |    |||| |||| ||||| ||| | |||| ||||||  | ||
Db    301 AAGGCTAAGATTGAGCTTTGGGAGAGGAACCAAAGGCATTATCTGGGAGAAGATTTAGAA 360

Qy    455 TCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGC 514
          ||  | ||| | ||||||||  || |   ||||||| ||||| ||  |  || | ||
Db    361 TCAATCAGCATAAAGGAGCTACAGAATCTGGAGCAGCAGCTTGACACTTCTCTTAAACAT 420

Qy    515 ATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGAT 574
          ||| | || ||||| || |||   |||  | | ||| |       |||||  |  |||||
Db    421 ATTCGCTCCAGAAAAAATCAACTAATGCACTAGTCCCTCAACCACCTCCAAAGAAAGGAG 480

Qy    575 AAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAA 634
          ||||   | |  ||  | ||||    ||||  |||| |||| | ||||||||||  |
Db    481 AAAGAAATACTGGAGGAAAACAGCATGCTTGCCAAACAGATAAAGGAGAGGGAGAGTATC 540

Qy    635 ACGGGTCAGCAAGAAGGACAATTAGTCCA 663
            |    ||  ||   |||| ||  ||
Db    541 CTAAGGACACATCAAAACCAATCAGAGCA 569
```

RESULT 12
US-09-853-450-11
; Sequence 11, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Martin F.
; APPLICANT: Pelaz, Soraya
; APPLICANT: Ditta, Gary
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
; TITLE OF INVENTION: Exhibiting Modulated Reproductive Development
; FILE REFERENCE: 19452A-002400US
; CURRENT APPLICATION NUMBER: US/09/853,450
; CURRENT FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Brassica oleracea
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(756)
; OTHER INFORMATION: CAULIFLOWER
US-09-853-450-11

Query Match 23.3%; Score 247.8; DB 9; Length 756;
Best Local Similarity 66.1%; Pred. No. 5e-55;
Matches 376; Conservative 0; Mismatches 187; Indels 6; Gaps 1;

```
Qy    101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
          |||||||| |||||||||  |  ||||||||||||||||||||||||  || |||||  ||
Db      1 ATGGGAAGGGGTAGGGTTGAAATGAAGAGGATAGAGAACAAGATCAACCGACAAGTGACG 60

Qy    161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
          || || || |||||  ||||| |  | ||||||| ||||||||||  | || || |||
Db     61 TTTTCGAAAAGAAGAGCTGGTCTTTTGAAGAAAGCCCATGAGATCTCGATCCTTTGTGAT 120

Qy    221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
          |||||||||  | || || ||||||||    || || |||||  ||||| || ||  | ||
Db    121 GCTGAGGTTTCCCTTATTGTCTTCTCCCATAAGGGGAAACTGTTCGAGTACTCGTCTGAA 180

Qy    281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
          ||||||||||||| | |||| |||||||| ||  | || |  ||  | || |||||  ||
Db    181 TCTTGCATGGAGAAGGTACTAGAACACTACGAGAGGTACTCTTACGCCGAGAAACAGCTA 240

Qy    341 GTTGGCCGAGACGTTTCACAAAGTGAA------AATTGGGTTCTAGAACATGCTAAGCTC 394
             |  | ||||  |       | || |       || |||     | ||| ||   | |||
Db    241 AAAGTTCCAGACTCTCACGTCAATGCACAAACGAACTGGTCAGTGGAATATAGCAGGCTT 300

Qy    395 AAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGAT 454
          |||||  |   |||||  |    |||| |||| ||||| ||| | |||| ||||||  | ||
Db    301 AAGGCTAAGATTGAGCTTTTGGAGAGAAACCAAAGGCATTATCTGGGCGAAGATTTAGAA 360

Qy    455 TCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGC 514
          ||  | ||| | ||||||||  || |   ||||||| ||||| ||  |  || | ||
Db    361 TCAATCAGCATAAAGGAGCTACAGAATCTGGAGCAGCAGCTTGACACTTCTCTTAAACAT 420

Qy    515 ATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGAT 574
          ||| | || ||||| || |||   |||  | | ||| |       |||||  |  |||||
Db    421 ATTCGCTCGAGAAAAAATCAACTAATGCACGAGTCCCTCAACCACCTCCAAAGAAAGGAG 480

Qy    575 AAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAA 634
          ||||   | |  ||  | ||||    ||||  |||| |||| | ||||||||||  |
Db    481 AAAGAAATACTGGAGGAAAACAGCATGCTTGCCAAACAGATAAGGGAGAGGGAGAGTATC 540

Qy    635 ACGGGTCAGCAAGAAGGACAATTAGTCCA 663
            |    ||  ||   |||| ||  ||
Db    541 CTAAGGACACATCAAAACCAATCAGAGCA 569
```

RESULT 13
US-09-938-842A-3474
; Sequence 3474, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3474
; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3474

Query Match 21.7%; Score 230; DB 9; Length 1241;
Best Local Similarity 100.0%; Pred. No. 2.7e-50;
Matches 230; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy    830 AACTATCTCACTCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACA 889
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      1 AACTATCTCACTCTTTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACA 60

Qy    890 TTCAGCATTTTTTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCAT 949
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TTCAGCATTTTTTTGGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCAT 120

Qy    950 ATTATATGTATGATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTC 1009
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 ATTATATGTATGATGGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTC 180

Qy   1010 ACTGCGTCTTAAGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAA 1059
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ACTGCGTCTTAAGAACAAAGATTCATATCTTGGTAATGATTTCTCATGAA 230


RESULT 14
US-09-853-450-27
; Sequence 27, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
;   LENGTH: 747
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(747)
;   OTHER INFORMATION: SEPALLATA1 (SEP1)
US-09-853-450-27

  Query Match             18.3%;  Score 194.4;  DB 9;  Length 747;
  Best Local Similarity   59.4%;  Pred. No. 3.8e-41;
  Matches  330;  Conservative    0;  Mismatches  226;  Indels    0;  Gaps    0;

Qy    101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
          ||||||||||| || ||  |||||||||||||||||||| ||||| || ||||| ||
Db      1 ATGGGAAGAGGAAGAGTAGAGCTGAAGAGGATAGAGAACAAAATCAACAGACAAGTAACG 60

Qy    161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
          ||  ||||| | |||   |||||| | |||||||| |||| | ||||||||||| |||
Db     61 TTTGCAAAGCGTAGGAACGGTTTGTTGAAGAAAGCTTATGAATTGTCTGTTCTCTGTGAT 120

Qy    221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
          ||||| ||||||||||||| |||||||   |   || || ||||  || | || || | |
Db    121 GCTGAAGTTGCTCTCATCATCTTCTCCAACCGTGGAAAGCTCTATGAGTTTTGCAGCTCC 180

Qy    281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
          ||   ||||   | || ||||||| || ||  |    |     ||| |  ||   || | |
Db    181 TCAAACATGCTCAAGACACTTGATCGGTACCAGAAATGCAGCTATGGATCCATTGAAGTC 240

Qy    341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
              |  |   |    | ||  ||| ||  |      |||| ||   ||||| ||||
Db    241 AACAACAAACCTGCCAAAGAACTTGAGAACAGCTACAGAGAATATCTGAAGCTTAAGGGT 300

Qy    401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
          |||  ||||   ||| |     |  | || ||| || | ||||| ||| | |   | |||
Db    301 AGATATGAGAACCTTCAACGTCAACAGAGAAATCTTCTTGGGGAGGATTTAGGACCTTTG 360

Qy    461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
          |  |  ||||||  |  |     | |||| ||| || || |   || |||||    || ||
Db    361 AATTCAAAGGAGTTAGAGCAGCTTGAGCGTCAACTGGACGGCTCTCTCAAGCAAGTTCGG 420

Qy    521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
          || |  ||||  ||    ||| | ||     | || |  || || || || ||  ||
Db    421 TCCATCAAGACACAGTACATGCTTGACCAGCTCTCGGATCTTCAAAATAAAGAGCAAATG 480

Qy    581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGT 640
          ||||  ||   ||| |   |  |    |  ||| |  | || | |       |   ||
Db    481 TTGCTTGAAACCAATAGAGCTTTGGCAATGAAGCTGGATGATATGATTGGTGTGAGAAGT 540

Qy    641 CAGCAAGAAGGACAAT 656
          || ||    |||  ||
Db    541 CATCATATGGGAGGAT 556


RESULT 15
US-09-853-450-29
; Sequence 29, Application US/09853450
; Publication No. US20020194645A1
; GENERAL INFORMATION:
;  APPLICANT: Yanofsky, Martin F.
;  APPLICANT:  Pelaz, Soraya
;  APPLICANT:  Ditta, Gary
;  APPLICANT:  The Regents of the University of California
;  TITLE OF INVENTION: Combinations of Genes for Producing Seed Plants
;  TITLE OF INVENTION:  Exhibiting Modulated Reproductive Development
;  FILE REFERENCE: 19452A-002400US
;  CURRENT APPLICATION NUMBER: US/09/853,450
;  CURRENT FILING DATE:  2001-05-09
;  NUMBER OF SEQ ID NOS: 61
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 29
;   LENGTH: 753
;   TYPE: DNA
;   ORGANISM: Arabidopsis thaliana
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(753)
;   OTHER INFORMATION: SEPALLATA2 (SEP2)
US-09-853-450-29

  Query Match             17.5%;  Score 185.8;  DB 9;  Length 753;
  Best Local Similarity   59.6%;  Pred. No. 6.6e-39;
  Matches  313;  Conservative    0;  Mismatches  212;  Indels    0;  Gaps    0;

Qy    101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
          ||||||||||| || ||  |||| ||||||||||||||||||| ||||| || ||||| ||
Db      1 ATGGGAAGAGGAAGAGTAGAGCTCAAGAGGATAGAGAACAAAATCAACAGACAAGTGACG 60

Qy    161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
          ||  | ||  | ||   |||||||||| || |||||| |||||  | ||||||||||||||
Db     61 TTTGCTAAACGTAGAAATGGTTTGCTGAAAAAAGCTTATGAGCTTTCTGTTCTCTGCGAT 120

Qy    221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
          ||||| ||  |||||||||||||||||    |   ||||| |||| ||| |  | || |  |
Db    121 GCTGAAGTCTCTCTCATCGTCTTCTCCAACCGTGGCAAGCTCTACGAGTTCTGCAGCACC 180

Qy    281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
          ||   ||||   | || ||| ||| | |||  |    | |   |||     ||   || | |
Db    181 TCCAACATGCTCAAGACACTGGAAAGGTATCAGAAGTGTAGCTATGGCTCCATTGAAGTC 240

Qy    341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
              |  |   | |  | |   ||| ||  |      |||  |    ||||| || |
Db    241 AACAACAAACCTGCTAAAGAGCTTGAGAACAGCTACAGAGAGTACTTGAAGCTGAAAGGT 300

Qy    401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
          |||  |||    ||  |     |  | || ||| || | || |||||||   |  ||
Db    301 AGATATGAAAATCTGCAACGTCAGCAGAGAAATCTTCTTGGAGAGGATCTTGGACCTCTG 360

Qy    461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
          |  |  ||||||||  |     | |||| ||| || || |   || | |||    || |
Db    361 AATTCAAAGGAGCTAGAGCAGCTTGAGCGTCAACTAGACGGCTCTCTGAAGCAAGTTCGC 420

Qy    521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
          |  |  ||||  ||   |||| | ||     | ||||  || ||    |||||  |   |
```

```
Db    421 TGCATCAAGACACAGTATATGCTTGACCAGCTCTCTGATCTTCAAGGTAAGGAGCATATC 480

Qy    581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGG 625
          ||||  |||  |||||   |  |    |  ||| |  | || | |
Db    481 TTGCTTGATGCCAACAGAGCTTTGTCAATGAAGCTGGAAGATATG 525
```

Search completed: February 2, 2003, 05:41:08
Job time : 98 secs

GenCore version 5.1.3

OM nucleic - nucleic search, using sw model

Run on: February 2, 2003, 05:41:13 ; Search time 2182 Seconds
(without alignments)
7882.496 Million cell updates/sec

Title: US-09-981-087A-1
Perfect score: 1062
Sequence: 1 cccagagagacataagaaag.....taatgatttctcatgaaata 1062

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	682.4	64.3	685	10	AV822813	AV822813 AV822813
c 2	435	41.0	455	10	AV794470	AV794470 AV794470
c 3	358.8	33.8	362	10	AV563010	AV563010 AV563010
c 4	350	33.0	396	10	AV817219	AV817219 AV817219
5	327.4	30.8	680	14	BQ148488	BQ148488 NF068H07F
6	327.2	30.8	951	12	BG838400	BG838400 Gc01_02c0
7	316	29.8	777	10	BE034098	BE034098 MG05B05 M
8	313.2	29.5	637	10	AW442341	AW442341 EST311737
9	313.2	29.5	729	10	AW442282	AW442282 EST311678
10	310.2	29.2	610	10	AW441624	AW441624 EST311020
11	310.2	29.2	640	13	BI930479	BI930479 EST550368
12	310.2	29.2	682	9	AI486645	AI486645 EST244966
13	308.8	29.1	653	10	AW930971	AW930971 EST356814
14	308.6	29.1	638	13	BM411808	BM411808 EST586124
15	305	28.7	634	10	AW441346	AW441346 EST310742
16	303.8	28.6	588	10	BE431620	BE431620 EST336435
17	302.8	28.5	639	10	AW223473	AW223473 EST300284
18	302	28.4	728	13	BM412293	BM412293 EST586620
19	301.2	28.4	535	9	AI489508	AI489508 EST247847
20	301	28.3	559	10	BE460489	BE460489 EST411908
21	301	28.3	618	13	BM535331	BM535331 EST588353
22	301	28.3	763	13	BM412720	BM412720 EST587058
23	297.8	28.0	560	10	AW222435	AW222435 EST299246
24	297.8	28.0	675	12	BG888499	BG888499 EST514350
25	297.8	28.0	782	14	BQ510995	BQ510995 EST618410
26	297.4	28.0	515	10	BE435803	BE435803 EST406881
27	296	27.9	528	10	AW220728	AW220728 EST297197
28	292.4	27.5	613	13	BM406322	BM406322 EST580649
29	291.8	27.5	534	9	AI898519	AI898519 EST267962
30	287.8	27.1	507	10	AW222490	AW222490 EST299301
31	286.8	27.0	514	10	AW930435	AW930435 EST340808
32	276.8	26.1	488	12	BF052183	BF052183 EST437430
33	275.6	26.0	755	14	BQ988395	BQ988395 QGF14M24.
34	274.6	25.9	654	14	BU010049	BU010049 QGJ12F07.
35	273	25.7	739	14	BQ986698	BQ986698 QGF10D09.
36	268.6	25.3	541	10	AW931475	AW931475 EST357318
37	268	25.2	548	10	AW932141	AW932141 EST357984
38	266.2	25.1	574	10	AW931861	AW931861 EST357704
39	264	24.9	638	9	AU236519	AU236519 AU236519
40	262.6	24.7	470	10	BE436562	BE436562 EST407640
41	262.6	24.7	495	10	AW254887	AW254887 ML1311 pe
42	262	24.7	586	14	BQ454007	BQ454007 sap03e02.
43	260.8	24.6	517	10	AW441991	AW441991 EST311387
44	260.4	24.5	581	9	AI487956	AI487956 EST246278
45	256.6	24.2	438	14	BQ149729	BQ149729 NF108A09F

ALIGNMENTS

RESULT 1
AV822813
LOCUS AV822813 685 bp mRNA linear EST 01-APR-2002
DEFINITION AV822813 RAFL5 Arabidopsis thaliana cDNA clone RAFL05-11-O24 5', mRNA sequence.
ACCESSION AV822813
VERSION AV822813.1 GI:19864865
KEYWORDS EST.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 685)
AUTHORS Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL Unpublished (2002)
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: mseki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998).This clone is in a

```
          modified pBluescript vector as a SstI/XhoI insert. Please visit our
          web site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for
          further details.
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL05-11-O24"
                     /clone_lib="RAFL5"
                     /dev_stage="rosette plants"
                     /lab_host="SOLR"
                     /note="Site_1: SstI; Site_2: XhoI; subjected to
                     dehydration-treated(1,2,5,10,24 hr)"
BASE COUNT      208 a    131 c    165 g    180 t      1 others
ORIGIN

  Query Match             64.3%;  Score 682.4;  DB 10;  Length 685;
  Best Local Similarity   99.7%;  Pred. No. 1.2e-160;
  Matches  683;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      39 TTTGGTCATTTCAGGGTTGTCGTTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAG 98
           || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TTGGGTCATTTCAGGGTTGTCGTTTCTCTCTCTTGTTCTTGAGATTTTGAAGAGAGAGAG 60

Qy      99 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 158
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 120

Qy     159 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 218
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 180

Qy     219 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 278
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 240

Qy     279 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 338
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 300

Qy     339 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 398
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 360

Qy     399 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 458
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 420

Qy     459 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 518
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 480

Qy     519 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 578
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 540

Qy     579 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 638
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 600

Qy     639 GTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAAT 698
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAAT 660

Qy     699 ACTGCGTAACCTCCTCCAGAGATGG 723
           ||||||||||||||| ||||||||
Db     661 ACTGCGTAACCTCCTNCAGAGATGG 685


RESULT 2
AV794470/c
LOCUS       AV794470                455 bp    mRNA    linear   EST 29-MAR-2002
DEFINITION  AV794470 RAFL8 Arabidopsis thaliana cDNA clone RAFL08-13-A09 3',
            mRNA sequence.
ACCESSION   AV794470
VERSION     AV794470.1  GI:19828453
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 455)
  AUTHORS   Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
            Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
            and Shinozaki,K.
  TITLE     Large scale analysis of Arabidopsis full-length cDNA (2002b)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
            al., submitted for publication) digested with BamHI and SalI. This
            clone is in a modified pBluescript vector. Please visit our web
            site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further
            details.
FEATURES             Location/Qualifiers
     source          1. .455
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL08-13-A09"
                     /clone_lib="RAFL8"
                     /dev_stage="rosette plants"
                     /lab_host="DH10B"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to
                     dehydration-treated (1, 2, 5, 10, 24 hr)"
BASE COUNT      150 a     87 c     85 g    133 t
ORIGIN

  Query Match             41.0%;  Score 435;  DB 10;  Length 455;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-98;
  Matches  435;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     625 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 684
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     455 GGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATGCTCCAACTCTTCTTCAGT 396

Qy     685 TCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGATGGCTTTGTGGAGAGAGTTGGGGG 744
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     395 TCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGATGGCTTTGTGGAGAGAGTTGGGGG 336

Qy     745 AGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTCTCTGCTTCCGGCTTGGATGTT 804
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     335 AGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTCTCTGCTTCCGGCTTGGATGTT 276

Qy     805 ACGTCCTACCACTACGAACGAGTAGAACTATCTCACTCTTTATAATATAATGATAATATA 864
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     275 ACGTCCTACCACTACGAACGAGTAGAACTATCTCACTCTTTATAATATAATGATAATATA 216

Qy     865 ATTAATGTTTAATATTTTCATAACATTCAGCATTTTTTTGGTGACTTATACTCATTATTA 924
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     215 ATTAATGTTTAATATTTTCATAACATTCAGCATTTTTTTGGTGACTTATACTCATTATTA 156

Qy     925 ATACCGATATGTTTTAGCTAGTCATATTATATGTATGATGGAACTCCGTTGTCGAGACGT 984
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     155 ATACCGATATGTTTTAGCTAGTCATATTATATGTATGATGGAACTCCGTTGTCGAGACGT 96

Qy     985 ATGTACGTAAGCTATCATTAGATTCACTGCGTCTTAAGAACAAAGATTCATATCTTGGTA 1044
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     95 ATGTACGTAAGCTATCATTAGATTCACTGCGTCTTAAGAACAAAGATTCATATCTTGGTA 36

Qy   1045 ATGATTTCTCATGAA 1059
          |||||||||||||||
Db     35 ATGATTTCTCATGAA 21


RESULT 3
AV563010/c
LOCUS       AV563010                362 bp    mRNA    linear   EST 07-SEP-2000
DEFINITION  AV563010 Arabidopsis thaliana green siliques Columbia Arabidopsis
            thaliana cDNA clone SQ179h11F 3', mRNA sequence.
ACCESSION   AV563010
VERSION     AV563010.1  GI:8734436
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 362)
  AUTHORS   Asamizu,E., Nakamura,Y., Sato,S. and Tabata,S.
  TITLE     A large scale analysis of cDNA in Arabidopsis thaliana: Generation
            of 12,028 non-redundant expressed sequence tags from normalized and
            size-selected cDNA libraries
  JOURNAL   DNA Res. 7, 175-180 (2000)
  MEDLINE   20363093
COMMENT     Contact: Erika Asamizu
            The First Laboratory for Plant Gene Research
            Kazusa DNA Research Institute
            Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
            Email: asamizu@kazusa.or.jp, URL:http://www.kazusa.or.jp/en/plant/.
FEATURES             Location/Qualifiers
     source          1. .362
                     /organism="Arabidopsis thaliana"
                     /strain="Columbia"
                     /db_xref="taxon:3702"
                     /clone="SQ179h11F"
                     /clone_lib="Arabidopsis thaliana green siliques Columbia"
                     /tissue_type="green siliques"
                     /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
                     XhoI".
BASE COUNT      124 a     67 c     72 g     99 t
ORIGIN

  Query Match             33.8%;  Score 358.8;  DB 10;  Length 362;
  Best Local Similarity   99.4%;  Pred. No. 1.8e-79;
  Matches  360;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    664 ATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGATGG 723
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    362 ATGCTCCAACTCTTCTTCAGTTCTTCTGCCTCAATACTGCGTAACCTCCTCCAAAGATGG 303

Qy    724 CTTTGTGGAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTC 783
          ||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 CTTTGTGAAGAGAGTTGGGGGAGAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTC 243

Qy    784 TCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCACTCT 843
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 TCTGCTTCCGGCTTGGATGTTACGTCCTACCACTACGAACGAGTAGAACTATCTCACTCT 183

Qy    844 TTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTTTTT 903
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 TTATAATATAATGATAATATAATTAATGTTTAATATTTTCATAACATTCAGCATTTTTTT 123

Qy    904 GGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTATGAT 963
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 GGTGACTTATACTCATTATTAATACCGATATGTTTTAGCTAGTCATATTATATGTATGAT 63

Qy    964 GGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTAAGA 1023
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     62 GGAACTCCGTTGTCGAGACGTATGTACGTAAGCTATCATTAGATTCACTGCGTCTTAAGA 3

Qy   1024 AC 1025
          ||
Db      2 AC 1


RESULT 4
AV817219/c
LOCUS       AV817219                396 bp    mRNA    linear   EST 01-APR-2002
DEFINITION  AV817219 RAFL9 Arabidopsis thaliana cDNA clone RAFL09-93-J11 3',
            mRNA sequence.
ACCESSION   AV817219
VERSION     AV817219.1  GI:19859090
KEYWORDS    EST.
SOURCE      thale cress.
  ORGANISM  Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1  (bases 1 to 396)
  AUTHORS   Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
            Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y.,
            Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y.
            and Shinozaki,K.
  TITLE     Large scale analysis of Arabidopsis full-length cDNA (2002b)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Motoaki Seki
            Plant Functional Genomics Research Group
            RIKEN Genomic Sciences Center
            3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
            Tel: 81-298-36-4359
            Fax: 81-298-36-9060
            Email: mseki@rtc.riken.go.jp
            An Arabidopsis full-length cDNA library was constructed essentially
            as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
            and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
            al., submitted for publication) digested with BamHI and SalI. This
            clone is in a modified pBluescript vector. Please visit our web
            site (http://www.gsc.riken.go.jp/e/plant/index _e.html) for further
            details.
FEATURES             Location/Qualifiers
     source          1. .396
                     /organism="Arabidopsis thaliana"
                     /db_xref="taxon:3702"
                     /clone="RAFL09-93-J11"
                     /clone_lib="RAFL9"
                     /dev_stage="plants at various developmental stages from
                     germination to mature seeds"
                     /lab_host="DH10B"
                     /note="Site_1: BamHI; Site_2: SalI; subjected to
                     dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
                     hr) treatments"
BASE COUNT      141 a     72 c     71 g    112 t
ORIGIN

  Query Match             33.0%;  Score 350;  DB 10;  Length 396;
  Best Local Similarity   99.5%;  Pred. No. 3e-77;
  Matches  372;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy    686 CTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGATGGCTTTGTGGAGAGAGTTGGGGGA 745
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    396 CTTCTGCCTCAATACTGCGTAACCTCCTCCAGAGATGGCTTTGTGGAGAGAGTTGGGGGA 337

Qy    746 GAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTCTCTGCTTCCGGCTTGGATGTTA 805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    336 GAGAACGGTGGTGCATCGTCGTTGACGGAACCAAACTCTCTGCTTCCGGCTTGGATGTTA 277

Qy    806 CGTCCTACCACTACGAACGAGTAGAACTATCTCACTCTTTATAATATAATGATAATATAA 865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    276 CGTCCTACCACTACGAACGAGTAGAACTATCTCACTCTTTATAATATAATGATAATATAA 217

Qy    866 TTAATGTTTAATATTTTCATAACATTCAGCATTTTTTTGGTGACTTATACTCATTATTAA 925
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    216 TTAATGTTTAATATTTTCATAACATTCAGCA-TTTTTTGGTGACTTATACTCATTATTAA 158
```

```
Qy     926 TACCGATATGTTTTAGCTAGTCATATTATATGTATGATGGAACTCCGTTGTCGAGACGTA 985
           ||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||
Db     157 TACCGATATGTTTTAGCTAGTCATATTATATGTATGAT-GAACTCCGTTGTCGAGACGTA 99

Qy     986 TGTACGTAAGCTATCATTAGATTCACTGCGTCTTAAGAACAAAGATTCATATCTTGGTAA 1045
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      98 TGTACGTAAGCTATCATTAGATTCACTGCGTCTTAAGAACAAAGATTCATATCTTGGTAA 39

Qy    1046 TGATTTCTCATGAA 1059
           |||||||||||||
Db      38 TGATTTCTCATGAA 25
```

RESULT 5
BQ148488

```
LOCUS       BQ148488                680 bp    mRNA    linear   EST 24-APR-2002
DEFINITION  NF068H07FL1F1064 Developing flower Medicago truncatula cDNA clone
            NF068H07FL 5', mRNA sequence.
ACCESSION   BQ148488
VERSION     BQ148488.1  GI:20285547
KEYWORDS    EST.
SOURCE      barrel medic.
  ORGANISM  Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE   1  (bases 1 to 680)
  AUTHORS   Torres-Jerez,I., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J.,
            Flores,H.R., Inman,J.T., Weller,J.W. and May,G.D.
  TITLE     Expressed Sequence Tags from the Samuel Roberts Noble Foundation
            Medicago truncatula flower library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: May GD
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73402, USA
            Tel: 580 221 7391
            Fax: 580 221 7380
            Email: gdmay@noble.org
            Insert Length: 680   Std Error: 0.00
            Plate: 068  row: H  column: 07
            Seq primer: TCACACAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .680
                     /organism="Medicago truncatula"
                     /db_xref="taxon:3880"
                     /clone="NF068H07FL"
                     /clone_lib="Developing flower"
                     /tissue_type="Developing flowers"
                     /dev_stage="Developmentally pooled. Contains a mixture of
                     very young, developing, fully-opened flowers and flowers
                     in early transition into pods."
                     /note="Vector: Lambda Zap; cDNA was prepared from polyA+
                     enriched, pooled samples of equivalent  amounts of total
                     RNA from very young, developing, fully-opened flowers  and
                     flowers transitioning into pods. The cDNA was
                     directionally ligated  into the Uni-Zap XR vector
                     (Stratagene) and packaged using the Gigapack  III Gold
                     packaging extracts. Phagemids containing cDNA inserts were
                     in  vivo excised from the recombinant Uni-ZAP XR vector
                     using ExAssist helper  phage and the E. coli strain
                     XL1-Blue MRF' (Stratagene). Excised  plasmids were plated
                     using SOLR cells."
BASE COUNT      247 a    110 c    134 g    184 t      5 others
ORIGIN
```

```
  Query Match            30.8%;  Score 327.4;  DB 14;  Length 680;
  Best Local Similarity  75.6%;  Pred. No. 1.5e-71;
  Matches  403;  Conservative    0;  Mismatches  130;  Indels    0;  Gaps    0;

Qy     101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
           |||||| || || || || ||| |||||||||| |||||||||||||||| ||||| |||
Db     130 ATGGGGAGGGGAAGAGTGCAGTTGAAGAGGATTGAGAACAAGATCAATAGACAAGTCACT 189

Qy     161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
           || |||||||||||||||||||| | |||||||| || |||||||||||||| || |||
Db     190 TTTTCAAAGAGAAGGTCTGGTTTGTTGAAGAAAGCACAAGAGATCTCTGTTCTTTGTGAT 249

Qy     221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
           ||||| ||||||||||  | |||||| | |||||||| || || ||||| |||| |||
Db     250 GCTGAAGTTGCTCTCATTATTTTCTCTACTAAAGGCAAGCTTTTTGAATACTCCAGCGAT 309

Qy     281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
           |||| ||||| |  || |||||||| |||||  | | ||  |||   || |  ||||||
Db     310 CCTTGTATGGAAAAAATTCTTGAACGATATGAAAGATGTTCGTATATGGAGAGGCAACTT 369

Qy     341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
           |||    | ||    |||| || |||||||||||||||||| || || || ||||||||
Db     370 GTTACAAGTGAGCAGTCACCAAATGAAAATTGGGTTCTAGAGCACGCAAAACTCAAGGCT 429

Qy     401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
           ||  | || || ||| ||| ||  |||||||||||||||| |||||  | |||   ||
Db     430 AGGATGGAAGTGCTTCAGAGAAATCAAAGGAATTTTATGGGAGAAGACTTGGATGGTTTA 489

Qy     461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
           |  | || ||||| ||||| ||||| || || || || ||| ||||| ||||    |||||
Db     490 NGTCTCAAAGAGCTTCAAAGTTTGGAACAACAACTTGATTCAGCTCTCAAACAAATTAGA 549

Qy     521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
           ||| | |||||||||| |||||  ||||| || ||    || || || |||||||| |||
Db     550 TCACGGAAGAACCAAGTTATGTATGAATCTATTTCANAACTTCANAAAAAGGATAAGGCC 609

Qy     581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAA 633
           | ||||| || |||||   |||   ||| ||||| ||||| |  ||||||
Db     610 CTNCAAGAACATAACAACATGCTCACCAAGAAGATAAAGGAAAAAGAGAAGGA 662
```

RESULT 6
BG838400

```
LOCUS       BG838400                951 bp    mRNA    linear   EST 25-MAY-2001
DEFINITION  Gc01_02c02_A Gc01_AAFC_ECORC_cold_stressed_Glycine_clandestina
            Glycine clandestina cDNA clone Gc01_02c02, mRNA sequence.
ACCESSION   BG838400
VERSION     BG838400.1  GI:14204722
KEYWORDS    EST.
SOURCE      Glycine clandestina.
  ORGANISM  Glycine clandestina
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
            Glycine.
REFERENCE   1  (bases 1 to 951)
  AUTHORS   Singh,J.A., Farah,S., Chapados,J., Couroux,P., De Moors,A., Harris
            ,L.J., Hattori,J.I., Ouellet,T., Robert,L.S., Sprott,D. and Tinker
            ,N.A.
  TITLE     Expressed Sequence Tags from Cold-Stressed Glycine clandestina
            Seedlings
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Singh,J.A.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            KW Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A
            0C6, Canada
            Tel: (613) 759-1662
            Fax: (613) 759-1701
            Email: singhja@em.agr.ca.
FEATURES             Location/Qualifiers
     source          1. .951
                     /organism="Glycine clandestina"
                     /cultivar="1035"
                     /db_xref="taxon:45687"
                     /clone="Gc01_02c02"
                     /clone_lib="Gc01_AAFC_ECORC_cold_stressed_Glycine_clandest
                     ina"
```

/tissue_type="Leaves, stem"
/note="Vector: Bluescript SK+/XhoI-EcoRI; Site_1: EcoRI; Site_2: XhoI; Plants incubated at 2 degrees under 12 hours of light/day. Harvested after only 2-3 days of cold treatment. cDNA was prepared with the Uni-Zap cDNA kit from Stratagene. Eco RI adapters were linked followed by digest with Xho I/Eco RI and ligated to pBluescript."
BASE COUNT 313 a 179 c 214 g 245 t
ORIGIN

Query Match 30.8%; Score 327.2; DB 12; Length 951;
Best Local Similarity 72.3%; Pred. No. 1.8e-71;
Matches 425; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

```
Qy      77 TTGAGATTTTGAAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAG 136
           |||| | ||     |   | | ||||||| || || || ||  |  |||||||||| |||
Db      36 TTGATAATTATTGAAAGTATATATATGGGGAGGGGAAGAGTGGATTTGAAGAGGATCGAG 95

Qy     137 AACAAGATCAATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCT 196
           ||||||||||||||||||||  | ||||| ||||||||||||||||||||| ||||||||
Db      96 AACAAGATCAATAGGCAAGTGGCGTTCTCGAAGAGAAGGTCTGGTTTGCTTAAGAAAGCG 155

Qy     197 CATGAGATCTCTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGC 256
           |  |||||||||| || |||||||||| || ||||||||||||||||  | ||||||
Db     156 CGCGAGATCTCTGTGCTGTGCGATGCTGATGTGGCTCTCATCGTCTTCTCCACTAAAGGC 215

Qy     257 AAACTCTTCGAATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGC 316
           || || ||||| || |||| ||| |||| ||||| || || ||||||| |||||  |
Db     216 AAGCTTTTCGACTACTCCAACGAACCTTGTATGGAAAGAATTCTTGAACTGTATGAGAGG 275

Qy     317 TATTTATATTCAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTT 376
           ||||  ||| | || | |||||||| |||     ||     ||| || |||||| ||||||
Db     276 TATTCTTATGCGGAGAGACAACTTGCTGGAGATGATCAAGCACCAAATGAAAACTGGGTT 335

Qy     377 CTAGAACATGCTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTT 436
           | |||||  |  ||||| ||||| || || ||| |||| ||  ||  ||||||||||
Db     336 ATGGAACACGAAAAGCTTAAGGCTAGGGTGGAGTTACTACAGCGAAATCAAAGGAATTTT 395

Qy     437 ATGGGGGAAGATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTC 496
           |||||  |||||||| ||    || |   | |  |  || || || ||||| || || ||
Db     396 ATGGGAGAAGATCTGGACAGCTTAAATCTTAGAGGACTTCAGAGTTTGGAACAACAACTT 455

Qy     497 GATGCAGCTATCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCT 556
           |||  |  |||  ||||    ||||||  |||  |||||||||||| |||   ||||| || |||
Db     456 GATTCCGCTCTCAAACACATTAGATCACAAAAGAACCAAGCCATGAATGAATCTATTTCT 515

Qy     557 GCGCTCCAGAAGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCTCAAAAAGATT 616
           |||||  ||||| ||||||||  |  | |  || |||||||| | |||   ||| |||||
Db     516 GCGCTTCAGAAAAAGGATAAGTCACTCCGGGAACACAACAACTTGCTCTCCAAGAAGATA 575

Qy     617 AAGGAGAGGGAGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAA 664
           |||||  ||||||  |    |  |  |||||   | |  | | |||
Db     576 AAGGATAAAGAGAAGGAGCTCGCCCCACAAGAGCAAGAGGGACTGCAA 623
```

RESULT 7
BE034098
LOCUS BE034098 777 bp mRNA linear EST 07-JUN-2000
DEFINITION MG05B05 MG Mesembryanthemum crystallinum cDNA 5' similar to mads-box protein, mRNA sequence.
ACCESSION BE034098
VERSION BE034098.1 GI:8329107
KEYWORDS EST.
SOURCE common iceplant.
ORGANISM Mesembryanthemum crystallinum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Caryophyllidae; Caryophyllales; Aizoaceae; Mesembryanthemum.
REFERENCE 1 (bases 1 to 777)
AUTHORS Bohnert,H.J., Borchert,C., Brazille,S., Brooks,J., Eaton,M., Ferrea,H., Kawasaki,S., McCollough,A., Michalowski,C.B., Palacio,C., Scara,G., Wheeler,M. and Zepeda,G.R.
TITLE Functional Genomics of Plant Stress Tolerance
JOURNAL Unpublished (2000)
COMMENT Contact: Michalowski,C.B.
University of Arizona
Bio Sciences West room 513, Tucson, AZ 85721, USA
Tel: 520-621-7982
Fax: 520-621-1697
Email: cbm@u.arizona.edu
An open reading frame exists.
Insert Length: 1 Std Error: 0.00.
FEATURES Location/Qualifiers
source 1. .777
/organism="Mesembryanthemum crystallinum"
/db_xref="taxon:3544"
/clone_lib="MG"
/tissue_type="roots"
/dev_stage="5-6 weeks"
/note="grown in hydroponics, stress 400 mM NaCl (in 0.5 Hoagland's), 30 h stress"
BASE COUNT 250 a 155 c 198 g 172 t 2 others
ORIGIN

Query Match 29.8%; Score 316; DB 10; Length 777;
Best Local Similarity 71.6%; Pred. No. 1.1e-68;
Matches 415; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

```
Qy      87 GAAGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCA 146
           ||  | ||| | | ||||| || |||||||||||||||||||||||| ||||| |||| |
Db      77 GAGAAAAGAAAAAAATGGGGAGGGGTAGGGTTCAGCTGAAGAGGATCGAGAATAAGATTA 136

Qy     147 ATAGGCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCT 206
           |  | || ||||| ||||| ||| |  |  |||| || | |||||||| ||||||||||
Db     137 ACCGTCAGGTTACCTTCTCGAAGCGTCGAACTGGGCTGTTGAAGAAAGCGCATGAGATCT 196

Qy     207 CTGTTCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCG 266
           |||||||||||||||||||||||| ||||||||||||||  |||| ||||| ||||| |
Db     197 CTGTTCTCTGCGATGCTGAGGTTGCGCTCATCGTCTTCTCCACCAAGGGCAAGCTCTTTG 256

Qy     267 AATATTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATT 326
           | ||  |  | ||  |||||||||||| ||| || || || ||||||  | |||| ||||
Db     257 AGTACGCTTCTGATGCTTGCATGGAGAAGATTCTCGAGCGATATGAAAGATATTCATATG 316

Qy     327 CAGACAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATG 386
           |||| | |||| |   |   |  |   |          || ||  |||  | | || ||||
Db     317 CAGAAAGACAAATGACTACACCTGGTCCTGATTCCCATGTAAGCTGGACTTTGGAGCATG 376

Qy     387 CTAAGCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAG 446
           | |||||||||||  ||  | ||  | ||| |||| ||  ||||| |||  ||||| ||||
Db     377 CAAAGCTCAAGGCTAGGCTGGAAATTCTTCAGAAAAATCAAAGGCATTACATGGGAGAAG 436

Qy     447 ATCTTGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTA 506
           | |||| | |||||||  | ||||||||  || |  || ||||| ||||| |||   |||
Db     437 AACTTGGTACGTTGAGTCTAAAGGAGCTTCAGAATTTAGAGCAGCAGCTTGATAGTGCTC 496

Qy     507 TCAAGAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGA 566
           | ||    || || || | |||||| |||   ||||  ||||| || |||| ||| ||||
Db     497 TTAAACATATCAGATCCAAAAAGAATCAACTCATGTATGAATCAATTTCTGAGCTGCAGA 556

Qy     567 AGAAGGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGG 626
           |||||||  ||||| ||||| || || ||||| | |||    || ||| | ||||||| ||
Db     557 AGAAGGACAAAGCATTGCAGGAGCAAAACAACTTGCTAACAAAGAAGGTCAAGGAGAAGG 616

Qy     627 AGAAGAAAACGGGTCAGCAAGAAGGACAATTAGTCCAATG 666
           ||||| ||| ||  |||    | |  |||  ||  || ||
Db     617 AGAAGGAAAAGGCGCAGGCTCAGGCGCAACAAGCACAGTG 656
```

RESULT 8
AW442341
LOCUS AW442341 637 bp mRNA linear EST 18-MAY-2001
DEFINITION EST311737 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA clone cLEN22N3 5', mRNA sequence.
ACCESSION AW442341
VERSION AW442341.1 GI:6977592

```
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 637)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
            Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.
            , Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .637
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEN22N3"
                     /clone_lib="tomato fruit red ripe, TAMU"
                     /tissue_type="pericarp"
                     /dev_stage="red ripe (7-20 days post-breaker)"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni;  Fruit were tagged at the
                     breaker stage (first sign of lycopene accumulation on the
                     blossom end of the fruit) and harvested 7 days
                     post-breaker (fully red-ripe), 10 days post breaker, and
                     20 days post-breaker (over-ripe).  20 day fruit which
                     showed external or internal signs of pathogenesis were
                     discarded.  Fruit were cut in half and the seeds and
                     locules were discarded prior to freezing the pericarp."
BASE COUNT      208 a    120 c    150 g    159 t
ORIGIN

  Query Match             29.5%;  Score 313.2;  DB 10;  Length 637;
  Best Local Similarity   73.1%;  Pred. No. 5.4e-68;
  Matches  402;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy      99 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 158
           | |||||||||||| || || ||| ||||| | |||||||||||| || ||  | |||||||
Db       1 AAATGGGAAGAGGAAGAGTCCAGTTGAAGCGAATAGAGAACAAAATTAACCGTCAAGTTA 60

Qy     159 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 218
           | ||||| ||  |  | |||||||||||| |||||||| |||||||||||||| || || |
Db      61 CCTTCTCGAAACGTCGATCTGGTTTGCTGAAGAAAGCCCATGAGATCTCTGTGCTTTGTG 120

Qy     219 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 278
           |||||||||||| | | || || || ||| | ||||| |||||||| |||||| ||| ||
Db     121 ATGCTGAGGTTGGTTTGATTGTTTTTTCTACTAAAGGAAAACTCTTTGAATATGCCAACG 180

Qy     279 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 338
           | || ||||||||||||||||||||| | ||||| | || | || | | || ||||| |
Db     181 ATTCCTGCATGGAGAGGATACTTGAAAGATATGAAAGATACTCATTTGCTGAGAAACAGC 240

Qy     339 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 398
           |||||      ||   | |       | ||  |||   || |||||||| || || ||||
Db     241 TTGTTCCTACTGATCATACCTCCCCGGTAAGCTGGACCCTTGAACATGCAAAACTTAAGG 300

Qy     399 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 458
           | ||| |||||||  ||  ||| |||| ||| | ||| | |||| |||||| | || || |
Db     301 CCAGACTTGAGGTTCTGCAGAGGAACCAAAAGCATTATGTGGGAGAAGATTTGGAGTCCT 360

Qy     459 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 518
           | ||  ||||||| || || |   ||||||| ||||| ||| ||||| | ||   ||||
Db     361 TAAGTATGAAGGAACTTCAGAATCTGGAGCACCAGCTTGATTCAGCTCTTAAACACATTC 420

Qy     519 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 578
```

```
Db     421 GATCAAGAAAGAATCAATTGATGCATGAGTCCATTTCTGTGCTTCAAAAAAAGGACAGAG 480

Qy     579 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 638
           | ||||| || || |||||   ||||   || ||| | ||||||||||||||  |   ||
Db     481 CATTGCAGGAGCAAAACAACCAGCTTTCGAAGAAGGTGAAGGAGAGGGAGAAAGAGGTGG 540

Qy     639 GTCAGCAAGA 648
           |||||| |
Db     541 CACAGCAAAA 550


RESULT 9
AW442282
LOCUS       AW442282                729 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST311678 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA
            clone cLEN22M20 5', mRNA sequence.
ACCESSION   AW442282
VERSION     AW442282.1  GI:6977533
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 729)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
            Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.
            , Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .729
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEN22M20"
                     /clone_lib="tomato fruit red ripe, TAMU"
                     /tissue_type="pericarp"
                     /dev_stage="red ripe (7-20 days post-breaker)"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; supplier: Giovannoni;  Fruit were tagged at the
                     breaker stage (first sign of lycopene accumulation on the
                     blossom end of the fruit) and harvested 7 days
                     post-breaker (fully red-ripe), 10 days post breaker, and
                     20 days post-breaker (over-ripe).  20 day fruit which
                     showed external or internal signs of pathogenesis were
                     discarded.  Fruit were cut in half and the seeds and
                     locules were discarded prior to freezing the pericarp."
BASE COUNT      240 a    135 c    173 g    181 t
ORIGIN

  Query Match             29.5%;  Score 313.2;  DB 10;  Length 729;
  Best Local Similarity   73.1%;  Pred. No. 5.5e-68;
  Matches  402;  Conservative    0;  Mismatches  148;  Indels    0;  Gaps    0;

Qy      99 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 158
           | |||||||||||| || || ||| ||||| | |||||||||||| || ||  | |||||||
Db       2 AAATGGGAAGAGGAAGAGTCCAGTTGAAGCGAATAGAGAACAAAATTAACCGTCAAGTTA 61

Qy     159 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 218
           | ||||| ||  |  | |||||||||||| |||||||| |||||||||||||| || || |
Db      62 CCTTCTCGAAACGTCGATCTGGTTTGCTGAAGAAAGCCCATGAGATCTCTGTGCTTTGTG 121

Qy     219 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 278
           |||||||||||| | | || || || ||| | ||||| |||||||| |||||| ||| ||
```

```
Db    122 ATGCTGAGGTTGGTTTGATTGTTTTTTCTACTAAAGGAAAACTCTTTGAATATGCCAACG 181

Qy    279 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 338
          | || ||||||||||||||||||||| | |||||  | || | || | | || ||||| |
Db    182 ATTCCTGCATGGAGAGGATACTTGAAAGATATGAAAGATACTCATTTGCTGAGAAACAGC 241

Qy    339 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 398
          |||||      ||   | |       | ||  |||   || ||||||| || || ||||
Db    242 TTGTTCCTACTGATCATACCTCCCCGGTAAGCTGGACCCTTGAACATGCAAAACTTAAGG 301

Qy    399 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 458
          | ||| ||||||| ||  ||| |||| ||| | ||| | |||| |||||| | || || |
Db    302 CCAGACTTGAGGTTCTGCAGAGGAACCAAAAGCATTATGTGGGAGAAGATTTGGAGTCCT 361

Qy    459 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 518
          | ||  ||||||| || || |   ||||||| ||||| ||| |||| | ||   ||||
Db    362 TAAGTATGAAGGAACTTCAGAATCTGGAGCACCAGCTTGATTCAGCTCTTAAACACATTC 421

Qy    519 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 578
          | |||||||||||| |||   |||   || ||||| |||| ||| || || ||||| | ||
Db    422 GATCAAGAAAGAATCAATTGATGCATGAGTCCATTTCTGTGCTTCAAAAAAAGGACAGAG 481

Qy    579 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 638
          | ||||| || || |||||   ||||   || ||| | ||||||||||||||  |   ||
Db    482 CATTGCAGGAGCAAAACAACCAGCTTTCGAAGAAGGTGAAGGAGAGGGAGAAAGAGGTGG 541

Qy    639 GTCAGCAAGA 648
            |||||| |
Db    542 CACAGCAAAA 551
```

RESULT 10
AW441624
LOCUS AW441624 610 bp mRNA linear EST 18-MAY-2001
DEFINITION EST311020 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA clone cLEN17J10 5', mRNA sequence.
ACCESSION AW441624
VERSION AW441624.1 GI:6976875
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 610)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.

FEATURES Location/Qualifiers
source 1. .610
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEN17J10"
/clone_lib="tomato fruit red ripe, TAMU"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2: Xho1; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 217 a 101 c 139 g 153 t
ORIGIN

Query Match 29.2%; Score 310.2; DB 10; Length 610;
Best Local Similarity 72.4%; Pred. No. 3.1e-67;
Matches 402; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

```
Qy     91 AGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAG 150
          ||| |   ||||||| ||||| || || ||  ||||||| || |||||||| || ||| |
Db     14 AGAAATTAATATGGGTAGAGGAAGAGTACAATTGAAGAGAATTGAGAACAAAATTAATCG 73

Qy    151 GCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGT 210
          |||||||||||| |||||||| | |||||||||| || ||||||||||||||||||||
Db     74 TCAAGTTACTTTTTCAAAGAGGCGATCTGGTTTGCTTAAAAAAGCTCATGAGATCTCTGT 133

Qy    211 TCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATA 270
          || |||||||||||| ||||  ||||| || |||||  | ||||| |||||||| || ||
Db    134 GCTTTGCGATGCTGAAGTTGGACTCATTGTTTTCTCAACTAAAGGAAAACTCTTTGAGTA 193

Qy    271 TTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGA 330
          ||| || |||||||||||||| ||||| |||||| | |||||  | || | |||| | ||
Db    194 TTCTACTGACTCTTGCATGGAAAGGATTCTTGAAAGGTATGAAAGGTACTCATATGCTGA 253

Qy    331 CAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAA 390
          |  || ||| ||      ||  ||  |       |  | |||  | | ||||||||||
Db    254 AAGGCAGCTTAATGCTACTGATATTATAACCCCGGGTAGCTGGACTTTGGAACATGCTAA 313

Qy    391 GCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCT 450
          ||| ||||| ||| |||||||  |  | |  ||| ||| | ||| |   || |||||  |
Db    314 GCTTAAGGCCAGACTTGAGGTTTTGCAAAGAAACCAAAAGCATTATGCAGGAGAAGAGTT 373

Qy    451 TGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAA 510
          ||  | |||||  |||| ||||| || |   |||| || ||||||||| | ||| | ||
Db    374 GGACACATTGAGTATGAAAGAGCTTCAGAATCTGGAACACCAGCTCGATTCTGCTCTTAA 433

Qy    511 GAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAA 570
          |  |||| | || ||||||||||||   |||   |||||||| |||| ||| || |||||
Db    434 GCACATTCGATCTAGAAAGAACCAATTGATGCATGAATCCATTTCTGAGCTTCAAAAGAA 493

Qy    571 GGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAA 630
          ||| || || |||||||| || |||||    |||.  ||  || |||||| ||||||||
Db    494 GGACAAGGCATTGCAAGAACAAAACAACAATCTTTCAAAGCAGGTTAAGGAAAGGGAGAA 553

Qy    631 GAAAACGGGTCAGCA 645
          | | ||  || ||
Db    554 AGAGATGGCCCAACA 568
```

RESULT 11
BI930479
LOCUS BI930479 640 bp mRNA linear EST 18-OCT-2001
DEFINITION EST550368 tomato flower, 8 mm to preanthesis buds Lycopersicon esculentum cDNA clone cTOC16P17 5' end, mRNA sequence.
ACCESSION BI930479
VERSION BI930479.1 GI:16244951
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 640)
AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Karamycheva,S.A., Tsai,J., Utterback,T., Van Aken,S., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
TITLE Generation of ESTs from tomato flower tissue, buds 8 mm - preanthesis
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute

```
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .640
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cTOC16P17"
                     /clone_lib="tomato flower, 8 mm to preanthesis buds"
                     /tissue_type="flower"
                     /dev_stage="buds 8mm to preanthesis"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; supplier: Cornell University; sequencing: The
                     Institute for Genomic Research; Flower buds and flowers
                     were taken from greenhouse plants (4-8 wks old, TA496).
                     They were immediately frozen in liquid nitrogen and then
                     size-separated while remaining frozen."
BASE COUNT      219 a    113 c    141 g    167 t
ORIGIN

  Query Match             29.2%;  Score 310.2;  DB 13;  Length 640;
  Best Local Similarity   72.4%;  Pred. No. 3.1e-67;
  Matches  402;  Conservative    0;  Mismatches  153;  Indels    0;  Gaps    0;

Qy      91 AGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAG 150
           ||| |   ||||||| ||||| || || ||  ||||||| || |||||||| || ||| |
Db       8 AGAAATTAATATGGGTAGAGGAAGAGTACAATTGAAGAGAATTGAGAACAAAATTAATCG 67

Qy     151 GCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGT 210
           |||||||||||| ||||||||  | ||||||||||| || ||||||||||||||||||||
Db      68 TCAAGTTACTTTTTCAAAGAGGCGATCTGGTTTGCTTAAAAAAGCTCATGAGATCTCTGT 127

Qy     211 TCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATA 270
           || |||||||||| |||| ||||| || |||||  | ||||| |||||||| || ||
Db     128 GCTTTGCGATGCTGAAGTTGGACTCATTGTTTTCTCAACTAAAGGAAAACTCTTTGAGTA 187

Qy     271 TTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGA 330
           ||| || ||||||||||||||| ||||| |||||| | |||||  | || | |||| | ||
Db     188 TTCTACTGACTCTTGCATGGAAAGGATTCTTGAAAGGTATGAAAGGTACTCATATGCTGA 247

Qy     331 CAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAA 390
           |  || ||| ||      ||  ||  |     |  |||  | | ||||||||||
Db     248 AAGGCAGCTTAATGCTACTGATATTATAACCCCGGGTAGCTGGACTTTGGAACATGCTAA 307

Qy     391 GCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCT 450
           ||| ||||| ||| ||||||| |  | |  ||| ||| | ||| |   || ||||| |
Db     308 GCTTAAGGCCAGACTTGAGGTTTTGCAAAGAAACCAAAAGCATTATGCAGGAGAAGAGTT 367

Qy     451 TGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAA 510
           ||  | |||||  |||| ||||| || |   |||| || ||||||||| | ||| | ||
Db     368 GGACACATTGAGTATGAAAGAGCTTCAGAATCTGGAACACCAGCTCGATTCTGCTCTTAA 427

Qy     511 GAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAA 570
           |  |||| | || ||||||||||||   |||  |||||||| |||| ||| || ||||||
Db     428 GCACATTCGATCTAGAAAGAACCAATTGATGCATGAATCCATTTCTGAGCTTCAAAAGAA 487

Qy     571 GGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAA 630
           ||| || || |||||||| || |||||    |||   ||  || ||||||| |||||||
Db     488 GGACAAGGCATTGCAAGAACAAAACAACAATCTTTCAAAGCAGGTTAAGGAAAGGGAGAA 547

Qy     631 GAAAACGGGTCAGCA 645
           | | ||  || ||
Db     548 AGAGATGGCCCAACA 562


RESULT 12
AI486645
LOCUS       AI486645                682 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST244966 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
            cLED11N8, mRNA sequence.
ACCESSION   AI486645
VERSION     AI486645.1  GI:4382016
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 682)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.
            , Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
            ,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
            ,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato carpel tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES             Location/Qualifiers
     source          1. .682
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLED11N8"
                     /clone_lib="tomato ovary, TAMU"
                     /tissue_type="carpel"
                     /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
                     /lab_host="XL1-Blue MRF'"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLED - Tomato Carpel EST Library. OligodT-primed and
                     directionally cloned  cDNA in vector Lamda ZAP II with 5'
                     and 3' ends located at the EcoRI and  XhoI sites,
                     respectively."
BASE COUNT      252 a    117 c    142 g    171 t
ORIGIN

  Query Match             29.2%;  Score 310.2;  DB 9;  Length 682;
  Best Local Similarity   72.4%;  Pred. No. 3.1e-67;
  Matches  402;  Conservative    0;  Mismatches  153;  Indels    0;  Gaps    0;

Qy      91 AGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAG 150
           ||| |   ||||||| ||||| || || ||  ||||||| || |||||||| || ||| |
Db      82 AGAAATTAATATGGGTAGAGGAAGAGTACAATTGAAGAGAATTGAGAACAAAATTAATCG 141

Qy     151 GCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGT 210
           |||||||||||| ||||||||  | ||||||||||| || ||||||||||||||||||||
Db     142 TCAAGTTACTTTTTCAAAGAGGCGATCTGGTTTGCTTAAAAAAGCTCATGAGATCTCTGT 201

Qy     211 TCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATA 270
           || |||||||||| |||| ||||| || |||||  | ||||| |||||||| || ||
Db     202 GCTTTGCGATGCTGAAGTTGGACTCATTGTTTTCTCAACTAAAGGAAAACTCTTTGAGTA 261

Qy     271 TTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGA 330
           ||| || ||||||||||||||| ||||| |||||| | |||||  | || | |||| | ||
Db     262 TTCTACTGACTCTTGCATGGAAAGGATTCTTGAAAGGTATGAAAGGTACTCATATGCTGA 321

Qy     331 CAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAA 390
           |  || ||| ||      ||  ||  |     |  |||  | | ||||||||||
Db     322 AAGGCAGCTTAATGCTACTGATATTATAACCCCGGGTAGCTGGACTTTGGAACATGCTAA 381

Qy     391 GCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCT 450
           ||| ||||| ||| ||||||| |  | |  ||| ||| | ||| |   || ||||| |
Db     382 GCTTAAGGCCAGACTTGAGGTTTTGCAAAGAAACCAAAAGCATTATGCAGGAGAAGAGTT 441

Qy     451 TGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAA 510
           ||  | |||||  |||| ||||| || |   |||| || ||||||||| | ||| | ||
Db     442 GGACACATTGAGTATGAAAGAGCTTCAGAATCTGGAACACCAGCTCGATTCTGCTCTTAA 501

Qy     511 GAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAA 570
```

```
Db    502 GCACATTCGATCTAGAAAGAACCAATTGATGCATGAATCCATTTCTGAGCTTCAAAAGAA 561

Qy    571 GGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAA 630

Db    562 GGACAAGGCATTGCAAGAACAAAACAACAATCTTTCAAAGCAGGTTAAGGAAAGGGAGAA 621

Qy    631 GAAAACGGGTCAGCA 645

Db    622 AGAGATGGCCCAACA 636


RESULT 13
AW930971
LOCUS       AW930971                653 bp    mRNA    linear   EST 18-MAY-2001
DEFINITION  EST356814 tomato fruit mature green, TAMU Lycopersicon esculentum
            cDNA clone cLEF42E18 5', mRNA sequence.
ACCESSION   AW930971
VERSION     AW930971.1  GI:8106372
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 653)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
            Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.
            , Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue
  JOURNAL   Unpublished (1999)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            5 prime sequence.
FEATURES             Location/Qualifiers
     source          1. .653
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEF42E18"
                     /clone_lib="tomato fruit mature green, TAMU"
                     /tissue_type="fruit pericarp"
                     /dev_stage="mature green (3-5 days pre-ripening)"
                     /lab_host="SOLR"
                     /note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
                     Xho1; cLEF - Fruit were tagged at the 1cm stage and
                     harvested 3-5 days prior to  ripening. Fruit were cut in
                     half to verify the seeds were indeed  'immature' and the
                     seeds and locules were discarded prior to freezing the
                     pericarp"
BASE COUNT      247 a    107 c    133 g    166 t
ORIGIN

  Query Match           29.1%;  Score 308.8;  DB 10;  Length 653;
  Best Local Similarity   72.5%;  Pred. No. 6.9e-67;
  Matches  400;  Conservative    0;  Mismatches  152;  Indels    0;  Gaps    0;

Qy     91 AGAGAGAGATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAG 150

Db    102 AGAAATTAATATGGGTAGAGGAAGAGTACAATTGAAGAGAATTGAGAACAAAATTAATCG 161

Qy    151 GCAAGTTACTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGT 210

Db    162 TCAAGTTACTTTTTCAAAGAGGCGATCTGGTTTGCTTAAAAAAGCTCATGAGATCTCTGT 221

Qy    211 TCTCTGCGATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATA 270

Db    222 GCTTTGCGATGCTGAAGTTGGACTCATTGTTTTCTCAACTAAAGGAAAACTCTTTGAGTA 281

Qy    271 TTCCACCGACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGA 330

Db    282 TTCTACTGACTCTTGCATGGAAAGGATTCTTGAAAGGTATGAAAGGTACTCATATGCTGA 341

Qy    331 CAAACAACTTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAA 390

Db    342 AAGGCAGCTTAATGCTACTGATATTATAACCCCGGGTAGCTGGACTTTGGAACATGCTAA 401

Qy    391 GCTCAAGGCAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCT 450

Db    402 GCTTAAGGCCAGACTTGAGGTTTTGCAAAGAAACCAAAAGCATTATGCAGGAGAAGAGTT 461

Qy    451 TGATTCGTTGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAA 510

Db    462 GGACACATTGAGTATGAAAGAGCTTCAGAATCTGGAACACCAGCTCGATTCTGCTCTTAA 521

Qy    511 GAGCATTAGGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAA 570

Db    522 GCACATTCGATCTAGAAAGAACCAATTGATGCATGAATCCATTTCTGAGCTTCAAAAGAA 581

Qy    571 GGATAAAGCCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAA 630

Db    582 GGACAAGGCATTGCAAGAACAAAACAACAATCTTTCAAAGCAGGTTAAGGAAAGGGAGAA 641

Qy    631 GAAAACGGGTCA 642

Db    642 AGAGATGGCCCA 653


RESULT 14
BM411808
LOCUS       BM411808                638 bp    mRNA    linear   EST 22-JAN-2002
DEFINITION  EST586124 tomato breaker fruit Lycopersicon esculentum cDNA clone
            cLEG57P12 5' end, mRNA sequence.
ACCESSION   BM411808
VERSION     BM411808.1  GI:18263427
KEYWORDS    EST.
SOURCE      tomato.
  ORGANISM  Lycopersicon esculentum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1  (bases 1 to 638)
  AUTHORS   Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai
            ,J., Bougri,O., Kirkness,E., Utterback,T., Van Aken,S., Ronning
            ,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
  TITLE     Generation of ESTs from tomato fruit tissue, breaker stage (2002)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: CUGI
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Email: http://www.genome.clemson.edu/orders/index.html
            This clone is available through the Clemson University Genomics
            Institute
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .638
                     /organism="Lycopersicon esculentum"
                     /cultivar="TA496"
                     /db_xref="taxon:4081"
                     /clone="cLEG57P12"
                     /clone_lib="tomato breaker fruit"
                     /tissue_type="Pericarp"
                     /dev_stage="breaker"
                     /lab_host="SOLR"
                     /note="Vector: pBluescriptSKmCUadapt; Site_1: EcoR1;
                     Site_2: XhoI; supplier: Boyce Thompson Institute;
                     sequencing: The Institute for Genomic Research.  Fruit
                     were harvested at the breaker stage (first sign of
                     lycopene accumulation on the blossom end of fruit).  Fruit
                     were cut in half and the seeds and locules were discarded
```

prior to freezing the pericarp."
BASE COUNT 217 a 115 c 138 g 168 t
ORIGIN

Query Match 29.1%; Score 308.6; DB 13; Length 638;
Best Local Similarity 72.8%; Pred. No. 7.8e-67;
Matches 398; Conservative 0; Mismatches 149; Indels 0; Gaps 0;

```
Qy      99 ATATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTA 158
           ||||||| ||||| || || ||  ||||||| || |||||||| || ||| | |||||||
Db      11 ATATGGGTAGAGGAAGAGTACAATTGAAGAGAATTGAGAACAAAATTAATCGTCAAGTTA 70

Qy     159 CTTTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCG 218
           |||| ||||||||  | ||||||||||| || ||||||||||||||||||| || ||||
Db      71 CTTTTTCAAAGAGGCGATCTGGTTTGCTTAAAAAAGCTCATGAGATCTCTGTGCTTTGCG 130

Qy     219 ATGCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCG 278
           ||||||| ||||  ||||| || |||||  | ||||| |||||||| || ||||| || |
Db     131 ATGCTGAAGTTGGACTCATTGTTTTCTCAACTAAAGGAAAACTCTTTGAGTATTCTACTG 190

Qy     279 ACTCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAAC 338
           |||||||||||||| ||||| |||||| | |||||  | || | |||| | || |  || |
Db     191 ACTCTTGCATGGAAAGGATTCTTGAAAGGTATGAAAGGTACTCATATGCTGAAAGGCAGC 250

Qy     339 TTGTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGG 398
           ||  ||     ||  ||  |      |  |  |||  | | ||||||||||||| ||||
Db     251 TTAATGCTACTGATATTATAACCCCGGGTAGCTGGACTTTGGAACATGCTAAGCTTAAGG 310

Qy     399 CAAGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGT 458
           | ||| |||||||  |  | |  ||| ||| | ||| |   || |||||  | ||  | |
Db     311 CCAGACTTGAGGTTTTGCAAAGAAACCAAAAGCATTATGCAGGAGAAGAGTTGGACACAT 370

Qy     459 TGAGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTA 518
           ||||  |||| ||||| || |   |||| || |||||||| | ||| | |||  ||||
Db     371 TGAGTATGAAAGAGCTTCAGAATCTGGAACACCAGCTCGATTCTGCTCTTAAGCACATTC 430

Qy     519 GGTCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAG 578
           | || ||||||||||||   |||   |||||||| |||| ||| || |||||||| || |
Db     431 GATCTAGAAAGAACCAATTGATGCATGAATCCATTTCTGAGCTTCAAAAGAAGGACAAGG 490

Qy     579 CCTTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGG 638
           | ||||||||| || |||||    |||   ||  || ||||||| ||||||||  | | ||
Db     491 CATTGCAAGAACAAAACAACAATCTTTCAAAGCAGGTTAAGGAAAGGGAGAAAGAGATGG 550

Qy     639 GTCAGCA 645
             || ||
Db     551 CCCAACA 557
```

RESULT 15
AW441346
LOCUS AW441346 634 bp mRNA linear EST 18-MAY-2001
DEFINITION EST310742 tomato fruit red ripe, TAMU Lycopersicon esculentum cDNA clone cLEN15B17 5', mRNA sequence.
ACCESSION AW441346
VERSION AW441346.1 GI:6976597
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
REFERENCE 1 (bases 1 to 634)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
FEATURES Location/Qualifiers
source 1. .634
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEN15B17"
/clone_lib="tomato fruit red ripe, TAMU"
/tissue_type="pericarp"
/dev_stage="red ripe (7-20 days post-breaker)"
/note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2: Xho1; supplier: Giovannoni; Fruit were tagged at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit) and harvested 7 days post-breaker (fully red-ripe), 10 days post breaker, and 20 days post-breaker (over-ripe). 20 day fruit which showed external or internal signs of pathogenesis were discarded. Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."
BASE COUNT 214 a 115 c 141 g 164 t
ORIGIN

Query Match 28.7%; Score 305; DB 10; Length 634;
Best Local Similarity 72.5%; Pred. No. 6.2e-66;
Matches 395; Conservative 0; Mismatches 150; Indels 0; Gaps 0;

```
Qy     101 ATGGGAAGAGGTAGGGTTCAGCTGAAGAGGATAGAGAACAAGATCAATAGGCAAGTTACT 160
           ||||| ||||| || || ||  ||||||| || |||||||| || ||| | |||||||||
Db       1 ATGGGTAGAGGAAGAGTACAATTGAAGAGAATTGAGAACAAAATTAATCGTCAAGTTACT 60

Qy     161 TTCTCAAAGAGAAGGTCTGGTTTGCTCAAGAAAGCTCATGAGATCTCTGTTCTCTGCGAT 220
           || ||||||||  | ||||||||||| || ||||||||||||||||||| || ||||||
Db      61 TTTTCAAAGAGGCGATCTGGTTTGCTTAAAAAAGCTCATGAGATCTCTGTGCTTTGCGAT 120

Qy     221 GCTGAGGTTGCTCTCATCGTCTTCTCTTCCAAAGGCAAACTCTTCGAATATTCCACCGAC 280
           ||||| ||||  |||| || |||||  | ||||| |||||||| || ||||| || |||
Db     121 GCTGAAGTTGGACTCATTGTTTTCTCAACTAAAGGAAAACTCTTTGAGTATTCTACTGAC 180

Qy     281 TCTTGCATGGAGAGGATACTTGAACGCTATGATCGCTATTTATATTCAGACAAACAACTT 340
           |||||||||||| ||||| |||||| | |||||  | || | |||| | || |  || |||
Db     181 TCTTGCATGGAAAGGATTCTTGAAAGGTATGAAAGGTACTCATATGCTGAAAGGCAGCTT 240

Qy     341 GTTGGCCGAGACGTTTCACAAAGTGAAAATTGGGTTCTAGAACATGCTAAGCTCAAGGCA 400
             ||     ||  ||  |      |  |  |||  | | ||||||||||||| |||||
Db     241 AATGCTACTGATATTATAACCCCGGGTAGCTGGACTTTGGAACATGCTAAGCTTAAGGCC 300

Qy     401 AGAGTTGAGGTACTTGAGAAGAACAAAAGGAATTTTATGGGGGAAGATCTTGATTCGTTG 460
           ||| |||||||  |  | |  ||| ||| | ||| |   || |||||  | ||  | |||
Db     301 AGACTTGAGGTTTTGCAAAGAAACCAAAAGCATTATGCAGGAGAAGAGTTGGACACATTG 360

Qy     461 AGCTTGAAGGAGCTCCAAAGCTTGGAGCATCAGCTCGATGCAGCTATCAAGAGCATTAGG 520
           ||  |||| ||||| || |   |||| || |||||||| | ||| | |||  |||| |
Db     361 AGTATGAAAGAGCTTCAGAATCTGGAACACCAGCTCGATTCTGCTCTTAAGCACATTCGA 420

Qy     521 TCAAGAAAGAACCAAGCTATGTTCGAATCCATATCTGCGCTCCAGAAGAAGGATAAAGCC 580
           || ||||||||||||   |||   |||||||| |||| ||| || |||||||| || ||
Db     421 TCTAGAAAGAACCAATTGATGCATGAATCCATTTCTGAGCTTCAAAAGAAGGACAAGGCA 480

Qy     581 TTGCAAGATCACAACAATTCGCTTCTCAAAAAGATTAAGGAGAGGGAGAAGAAAACGGGT 640
           |||||||| || |||||    |||   ||  || |||| || |||||||  | | ||
Db     481 TTGCAAGAACAAAACAACAATCTTTCAAAGCAGGTTAACGAAAGGGAGAAAGAGATGGCC 540

Qy     641 CAGCA 645
           || ||
Db     541 CAACA 545
```

Search completed: February 2, 2003, 08:10:55
Job time : 2197 secs